A COLLEGE TEXT–BOOK

OF

PHYSICS

BY

ARTHUR L. KIMBALL, Ph.D.

LATE PROFESSOR OF PHYSICS
IN AMHERST COLLEGE

FOURTH EDITION

REVISED BY
ARTHUR L. KIMBALL, Jr

NEW YORK
HENRY HOLT AND COMPANY

PREFACE TO FOURTH EDITION

This revision is somewhat more complete than the previous one.

A few more sections have been added, and at the same time a number of sections of doubtful value have been dropped out.

The important subject of vibratory motion has been more fully treated. The entire subject of electrostatics and electricity has been presented from the viewpoint of electrons where reasonably possible. The articles on the conduction of electricity through gases and related material have been largely rewritten. Electric waves and their applications appear in the more logical position just before the chapter on light, and a few sections have been added on X-ray spectroscopy after the subject of diffraction in light.

An entire new chapter, entitled "Radiation and the Quantum Theory," has been added at the end of the book in slightly smaller type, as it is felt that a text-book of general college physics should contain some sort of introduction to "modern physics." This chapter is somewhat more difficult than the book proper, as it must be, to have real value. Most teachers will not care to use it in class work, but it will at least afford an opportunity for those interested to have a glimpse of what leading men in physics are now thinking about.

Acknowledgment is due particularly to Professor John Zeleny, of Yale University, for his careful reading of the manuscript and helpful suggestions.

<div style="text-align: right">A. L. K., Jr.</div>

Schenectady, N. Y.
May 8, 1929

PREFACE TO THIRD EDITION

In this revision an attempt has been made to conform as closely as possible to that already outlined and started by Professor Kimball before his death, so that the book shall remain his work, and be representative of his ideas.

The book is a few pages longer than before because, although additions have been made, it seems inadvisable for the present to omit or to change the original material except in a very few places.

The writer recognizes the difficulty of the task he has undertaken. He hopes that his knowledge of Professor Kimball's ideals and methods of teaching physics from many years of association with him, together with his own experience as a teacher, is a sufficient reason for attempting it.

Acknowledgment is due especially to Professor F. A. Saunders of Harvard University whose suggestions and criticisms of the manuscript have been most helpful. Acknowledgment should be made also to Professors J. O. Thompson, W. K. Green, P. I. Wold, W. S. Kimball, and to many others who have aided in the preparation of this revision.

A. L. K., Jr.

PREFACE TO REVISED EDITION

Recent advances in physical science having made it necessary to rewrite some paragraphs of the earlier edition, especially those relating to X-rays and the electron theory of matter, advantage has been taken of the opportunity to make a few additional changes which class-room experience has shown to be desirable. Certain paragraphs relating to force and motion, which had been introduced before the section on statics, are now placed among the introductory paragraphs to kinetics, where they fall in better with the logical development of the subject. The electro magnetic units, volt, ampere and ohm, are defined and introduced

earlier than before. The sections on wireless telegraphy have been made more complete and wireless telephony is touched upon. A section also has been added treating of the flicker photometer. At the end of the volume a short discussion of Carnot's cycle and the thermodynamic basis of the absolute scale of temperature has been introduced as an appendix, also a proof is given of Newton's wave formula. Quite a number of new problems have been added, but the old problems have been found to serve their purpose well and are for the most part retained.

The author gratefully acknowledges his indebtedness to Dr. G. S. Fulcher, and to Professors W. E. McElfresh and D. C. Miller, for valuable suggestions and criticisms.

<div align="right">A. L. K.</div>

AMHERST, MASS.
    July, 1917.

_____

PREFACE

In offering this work to my fellow teachers, a word of explanation is due. The book was undertaken some years ago when the writer felt the want of a text-book adapted to the needs of students taking the general first year course in college. As the work has slowly progressed several text-books of very similar aim have appeared, and it must be admitted that the call is not so imperative now as formerly; and yet it is hoped that the treatment here presented may meet some still existing demand and so justify its existence.

What may be called the physical rather than the mathematical method has been preferred in giving definitions and explanations, because it is believed that the ideas presented are more easily grasped and more tenaciously held when the mind forms for itself a sort of picture of the conditions, instead of merely associating them with the symbols of a formula.

There are many minds that do not easily grasp mathematical reasoning even of a simple sort; and it is often the case also that a student who may be able to follow an algebraic deduction step by

step has very little idea of the significance of the whole when he reaches the end. Algebra is not his native tongue and it takes considerable time and experience for him to learn to think in it. And while all will agree that for the more advanced study of physics, mathematics is quite indispensable, many will grant that in a general course, which is to furnish to most of those taking it all that they will ever know of physics as a science, the ideas and reasonings should be presented as directly as possible and in the most simple and familiar terms.

This then has been the central aim in the preparation of this book; to give the student clear and distinct conceptions of the various ideas and phenomena of physics, and to aid him in thinking through the relations between them, to the end that he may see something of the underlying unity of the subject; and to carry out this aim in such a manner that students may not be repelled by any unnecessary prominence of symbolic methods, and yet that the treatment may have all the exactness and precision in statement and deduction which the subject demands.

This is a large ambition and I cannot hope to have been wholly successful, but I shall be grateful if my attempt is found in any degree to have subserved its purpose.

My grateful acknowledgments are due to Dr. G. S. Fulcher of the University of Wisconsin, who has read nearly all the manuscript with great care, and to whom I am indebted for important suggestions, and to my colleague Professor J. O. Thompson whose criticism at all stages of the work and painstaking correction of the proof has been most helpful.

<div style="text-align:right">A. L. K.</div>

CONTENTS

INTRODUCTION

MECHANICS

I. General Principles.

II. Statics.

III. Kinetics of a Particle.

CONTENTS

ix

those circumstances without which the given event does not occur; and then we seek to determine the effect of each of these circumstances separately, and exhibit, if possible, each such effect as a special instance under some general law.

For example, the complex motion of a ball struck by a bat is found to be dependent on the motion given to it by the blow of the bat, on the presence of the earth, and on air resistance.

We first try and determine how a body moves when set free in the presence of the earth without any initial blow or impulse and in a vacuum. We find in this way an unvarying rule of motion that applies to all bodies of whatever size or shape, and we call it the law of falling bodies. That part of the motion of the ball which depends only on the nearness of the earth is but a special instance under this law. Now, making allowance for the motion due to the earth, we seek to determine that part of the motion due to the initial blow, and here again we find that the actual motion seems to be exactly according to a general rule which is found to hold whenever an impulsive force acts on a mass. And finally we investigate the effect of air resistance, determining how it affects a body at rest and how it modifies the motion of a body moving through it, and here again certain general rules are found which apply not only to the special case under consideration, but to all cases of bodies moving through air. When the effects of all three circumstances are taken into account, the motion is found to be exactly accounted for, and is then said to be explained.

Leverrier and Adams, in analyzing the motion of the planet Uranus, found that after taking account of all the known circumstances, such as the attractions of the sun and other planets upon it, there still remained a part of its motion which was not accounted for, and assuming it to be due to an unknown planet they computed its position and mass, and thus the planet Neptune was discovered.

But in analyzing our problem we may go deeper and show that the motion of the ball near the earth is such as would result from a force urging the two bodies together, and we may then discover that it is merely a special instance of the law that all bodies are influenced by forces urging them together or, in other words, that all bodies attract each other. When we can show

also that the forces between the air and the moving body are due to the motion given to the air and so are simply particular exhibitions of the general rule which holds whenever matter is set in motion, we feel that a still higher degree of understanding is reached.

By such a process all the complex facts of nature are assigned their places in an orderly system. But a limit is soon reached beyond which the mind cannot go, because thinking is conditioned by experience, and even in its profoundest theories and speculations the mind must employ those conceptions which it has obtained from the world about it.

3. Experiment. Physics is an experimental science, its generalizations rest solely upon experiment, and although reasoning upon established facts has often led to the discovery of new truths of great importance, the final appeal must always be to experiment. If the deduction is thus disproved, it appears either that the reasoning was wrong or that there are certain elements entering into the problem that were neglected. In seeking for the causes of such discrepancies new truths have often been discovered.

An experiment is a combination of circumstances brought about for the purpose of testing the truth of some deduction or for the discovery of new effects. The usual course of an experimental inquiry is to modify the circumstances one by one, noting the corresponding effect until the influence of each is thoroughly understood.

4. Necessary Assumptions. In every experimental science it is assumed that the same causes always produce the same effects and that the position of the event as a whole in either time or space only affects the absolute time and position of the result, provided there is no change in the *relative* time or space relations of the various circumstances involved. For example, if all the other circumstances are the same, a stone will fall in exactly the same manner next week as it does to-day, or if the solar system be changing its place in space no change in the manner of the stone's falling will take place from that cause alone.

Experience up to this time has justified these assumptions, and without them progress in physical science would be impossible.

5. Force. Our ideas of force are derived primarily from muscular effort. It requires an effort to lift a weight, to throw a ball, or to compress a spring. The upward pull upon a weight at any instant while it is being lifted is a force acting on it, and the downward tendency of the weight which the pull opposes is also a force.

Anything that serves to accomplish what would require muscular exertion to bring about, exerts a force. Thus a support exerts a force on the weight which rests upon it, and the weight exerts an equal and opposite force on the support, compressing it. A bat exerts a force against a ball in giving it motion, a clock spring exerts a force against the stop that prevents it from unwinding.

6. Matter. Through our muscular sense and sense of touch we are made conscious of bodies around us which resist compression, and may, therefore, be said to occupy space. Such bodies are said to be *material substances* or made of *matter*.

Every object that we know of possesses weight; that is, it requires some muscular effort to support it, or if it is hung on a spring the spring is stretched. What we call its weight is a force urging it toward the ground, and as the weight of two quarts of water is twice that of one quart we are led to think of the weight of a body as a proper measure of the quantity of matter which it contains.

But besides weight all bodies have inertia; that is, to produce a definite change per second in the motion of a body a certain force is required.

If a given body is isolated from other portions of matter, it may be heated or cooled or bent or twisted or compressed into small volume or allowed to expand into a large one, but in all these changes its weight and its inertia remain unchanged.

It will be seen later that if the body is taken from one place on the earth to another its *weight* also may change, so that *the only general property of a given portion of matter that cannot be changed is its inertia.*

It is this property, therefore, by which quantities of matter are defined, and two bodies which have equal inertias are

said to have equal *masses* or to contain equal quantities of matter.*

The actual comparison of two masses, however, is usually made by weighing, since under the ordinary circumstances of weighing, bodies which have equal weights have equal inertias.

7. Conservation of Matter. The mass of a given portion of matter as measured by its inertia cannot be changed by any process known to man. Not only may a piece of wood be bent and twisted or compressed without changing its mass, but it may be burned in the fire, and chemistry shows that if the ashes and vapors and gases that have come from it are collected and separated from the gases of the air with which they may have united, it will be found that the united mass of the ash and the gases and vapors is the same as the mass of the original piece of wood.

This principle is known as the *conservation of matter*, and is established by innumerable experiments, both physical and chemical.

8. States of Matter. Different kinds of matter differ greatly in the power of preserving their shape. Some, such as steel or copper, offer very great resistance to any attempt to change their forms. Such bodies are said to be *rigid* or *solid* bodies. Others, like water or air, have no permanent shape, but flow under the action of the weakest forces and take the shapes of the vessels containing them; they are called *fluids*. There are no substances that are either perfectly rigid or that are perfect fluids, for the most rigid bodies may be distorted, and those substances that flow most freely offer some resistance to change of form. In some cases it is difficult to say whether a substance is to be regarded as solid or fluid.

Fluids are again divided into *liquids* and *gases*. *Liquids* are those fluids that can have a free surface and do not change much in volume under great changes in pressure. A mass of liquid has a nearly definite bulk though no permanent shape. Water is an example of a liquid.

Gases, on the other hand, are fluids that do not have a free sur-

---

* According to Einstein's theory of relativity (§ 1009) mass increases with velocity of motion, but for all ordinary velocities this increase of mass is far too small to be detected by the most refined measurement.

face, but completely fill the containing vessel, however much it may be enlarged.

A mass of gas may be regarded as having neither permanent shape nor size, since both of these are entirely determined by the vessel which contains it. Air is a familar example of a gas.

Units and Measurements

9. Measurement. The exact measurement of all the quantities involved in any phenomenon is a very important part of its study. It is largely owing to the recognition of this that such great advances have been made in physics during the last two hundred years.

Every measurement is essentially a comparison. A quantity to be measured is compared with another quantity of the same kind called the *unit*. Thus to measure a length it is necessary to find how many times the unit of length is contained in the given length.

The unit must be of the same nature as the quantity which is to be measured, since only like things can be compared. There must, therefore, be as many different kinds of units as there are kinds of quantities to be measured.

10. Absolute Measurements. Each of the units employed might be arbitrarily chosen without reference to any other; the inch might be taken as the unit of length, the square foot as the unit of surface, and the quart as the unit of volume; but such a practice would lead to endless complications, especially when several different units are used in the same calculation, for it would be necessary in such a case to keep constantly in mind the number of square inches in a square foot, the number of cubic inches in a quart, etc. It is far simpler after choosing the inch as the unit of length to take the square inch as the unit of surface and the cubic inch as that of volume. The same principle applies in the case of all other units; none should be chosen arbitrarily which can be directly derived from those which have been previously selected.

A system such as this in which there are a few arbitrarily chosen fundamental units, between which no known connection exists and from which all other units are derived without

introducing any new arbitrary factors, is known as an *absolute* system of units.

11. Fundamental Units. All the phenomena of nature are manifested to us in *time* and in *space*, through the agency of *matter*. It is natural, then, that the fundamental units adopted as the basis of the system of measurement used in physics, should be the units of *time, length,* and *mass*. These are also convenient units, for lengths, times, and masses may be compared with great ease and precision, and all units that relate only to mass, motion and force or that depend on these by definition may be derived directly from them.

Physicists usually employ what is called the *centimeter-gram-second* (C. G. S.) system of absolute units in which the centimeter, gram, and second are taken as the units of length, mass, and time, respectively.

This system has the advantage of being in use by physicists all over the world, and therefore results expressed in its units are intelligible everywhere. But an absolute system might be based on any three units of length, mass, and time whatever. Thus a *foot-pound-second* system is used extensively in English-speaking countries.

12. Unit of Length. The unit of length in the C. G. S. system of absolute measurement is the *centimeter*, or one-hundredth part of a meter. The meter is the distance between the ends of a bar of platinum which is kept in Paris and known as the *Mètre des Archives*, the bar being measured when at the temperature of melting ice. This bar was constructed by Borda for the French Government, and was adopted by them in 1799 with the view to its becoming a universal standard of length; it was intended to be exactly one ten-millionth part of the distance from the equator to the pole measured along a meridian on the earth.

It is now known that the earth's quadrant is about 10,000,856 meters in length, but as distances can be more easily and accurately compared with the length of the bar at Paris than with the length of the earth's quadrant, the former still continues to be the standard of length.

The English standard of length from which the foot and inch are determined is the standard yard, which is the " distance between the centers of the transverse lines in the two gold plugs

in the bronze bar deposited in the office of the Exchequer" measured at the temperature of 62° F. This standard yard represents about the average length of the early yard measures that were in use, which were probably adopted as being half the distance which a man can stretch with his arms.

$$
\begin{array}{lcl}
1 \text{ yard} & = & 91.43835 \text{ centimeters.} \\
1 \text{ foot} & = & 30.47945 \text{ centimeters.} \\
1 \text{ inch} & = & 2.54 \text{ centimeters.} \\
1 \text{ mile} & = & 1.60935 \text{ kilometers.} \\
1 \text{ meter} & = & 39.37 \text{ inches.}
\end{array}
$$

13. Unit of Mass. The unit of mass in the C. G. S. system is the *gram*, or one-thousandth part of the standard kilogram, which is a mass of platinum kept at Paris and known as the *Kilogramme des Archives.* The standard kilogram was intended to represent the exact mass of a cubic decimeter of distilled water at its greatest density or at the temperature 4° C.

The gram is, therefore, equal to the mass of a cubic centimeter of pure water at 4° C. This relation between the cubic centimeter and the gram is exceedingly convenient, for it enables us to determine the volume of an irregular vessel from the weight of water which it can contain. But it is not a direct relation like that between the unit of length and unit of volume. Aside from convenience, there is no reason why a cubic centimeter of copper or mercury or of anything else might not have been taken as the unit of mass.

Since two masses may be compared with a far higher degree of accuracy than that with which the weight of a cubic centimeter of water can be determined, the *Kilogramme des Archives is the real standard on which all metric weights are based.*

The unit of mass in common use in engineering is the pound = 453.59 grams.

14. Unit of Time. *Intervals of time are always compared by the motions of bodies.* Two intervals of time are *defined as* equal when a body, moving under exactly the same circumstances in both cases, moves as far in the one time as in the other. The heavenly bodies have in their motions always furnished measures of time. One of the simplest natural units of time is the period of rotation of the earth, which is the interval of time between two

successive meridian passages of the same star. This is known as the *sidereal* day, and time reckoned in this way is called sidereal time. By considering the possible effect of tidal friction in retarding the earth's motion, Adams concludes that the period of rotation of the earth has not changed by more than one-thirtieth of a second in 3000 years.

The ordinary day is determined not by the rising and setting of the stars, but by the motion of the sun. When the sun is on the meridian it is said to be *solar* or *apparent* noon. The interval of time between two successive apparent noons is called the apparent or solar day. It is this time which is indicated by the sun dial. By means of clocks, which are machines constructed to run with great uniformity, one solar day may be compared with another, and it is thus found that they are not of equal length. The average length of the solar days in a year is known as the *mean solar day*.

The ordinary standard time used in everyday life is mean solar time.

The unit of time in the C. G. S. system is the mean solar second or the 86400th part of a mean solar day.

MECHANICS

I. GENERAL PRINCIPLES

15. Definitions. *Mechanics* treats of the motions of masses and of the effect of forces in causing or modifying those motions. It includes those cases where forces cause relative motions of the different parts of an elastic body causing it to change its shape or size, as when a gas is compressed or a spring bent. Such changes in size or shape of different portions of a body are called *strains*. Bodies which do not suffer strain when acted on by forces are said to be *rigid*.

All known bodies yield more or less to distorting or compressing forces, but when considering the motion of a body as a whole, all bodies in which the strains are small may be regarded as practically rigid. Thus we may treat the motion of a grindstone or of a shell from a rifled gun as though these bodies were rigid, though we know that they are slightly strained by the forces acting.

Mechanics is usually subdivided into kinematics and dynamics.

Kinematics treats of the characteristics of different kinds of motion, and of the modes of strain in elastic bodies without reference to the forces involved.

Dynamics treats of the effect of forces in causing or modifying the motions of masses and in producing strains in elastic bodies. It is usual to treat *dynamics* under the heads *statics* and *kinetics*.

Statics is that part of dynamics which deals with bodies in equilibrium or when the several forces that may be involved are so related as to balance or neutralize each other, so far as giving motion to the body as a whole is concerned.

Kinetics is that part of dynamics which treats of the effect of forces in changing the motions of bodies.

Ideas and Definitions of Kinematics

16. Motion Relative. When a body is changing its position it is said to be in motion. There is no way of fixing the posi-

tion of a body except by its distance from surrounding objects. When it is said, therefore, that a body has moved, it is always meant that there has been a change in its position with reference to some other objects regarded as fixed, or in other words, there has been *relative* motion. Thus we know only relative motion, and when we speak of an object as at rest we usually mean with reference to that part of the earth's surface in our vicinity.

17. Displacement. The distance in a straight line from one position of the body to another is called its displacement from the first position. To completely describe any displacement, its amount and direction must both be given.

If an extended rigid body is displaced, as when a book is moved on a table, it may be moved in such a way that its edges will remain parallel to their original directions, in which case the displacements of all points in the body will be the same both in amount and direction. The motion is said to be one of simple *translation* without *rotation*. But in general when a rigid body is moved there is rotation as well as translation, so that to bring it into the second position from the first we may first imagine it to be translated till some point in the object is brought into its second position. Then by a rotation about a suitable axis through that point the whole body may be brought into the second position.

18. Vectors and Their Representation. All quantities which involve the idea of direction as well as amount are said to be *vector quantities* or *vectors*. Such are displacements, velocities, forces, etc. While quantities having magnitude only, without any reference to direction, are known as *scalar* quantities. Volume, density, mass, and energy are scalar magnitudes. A vector quantity is represented by a straight line which indicates by its direction the direction of the vector, and by its length the magnitude of the vector, the length being measured in any convenient units, provided the same scale is used throughout any one diagram or construction.

It must be remembered, however, that a vector represented by a line AB is not the same as that represented by BA, one is the opposite of the other, or $AB = -BA$. This will be evident if AB represents a displacement from A to B. A displacement BA

will exactly undo what the other accomplished, and bring the body back to its starting point. The straight line representing a vector is, therefore, commonly represented with an arrow head indicating its positive direction.

19. Composition of Displacements. If a man in a railway car were to go directly across from one side to the other, say from A to B (Fig. 1), then the line AB will represent both in amount and direction his displacement considered only with respect to the car. But if the car is in motion and in the meantime has advanced through the distance AC, the man will evidently come to D instead
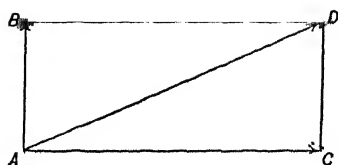
Fig. 1

of to B. The displacement of the car with reference to the earth is AC and the displacement of the man relative to the earth is AD. This is called the *resultant* displacement of the man, of which AB and AC are the components.

Another way of stating this is that the man received simultaneously two displacements AB and AC, for if he had not been displaced in the direction AC he would have gone to B, while if he had not had the displacement AB he would have been carried to C.

From the above it is evident that the resultant of any number of simultaneous displacements may be found just as if they had been taken successively.

For example, let it be required to find the resultant of four displacements represented in amount and direction by the vectors A, B, C, D (Fig. 2). If A were the only displacement, the body would be brought from O to a, but B is also a component displacement, therefore draw B' equal and parallel to B, and the result of the two displacements will be represented by the distance Ob. Then in like manner draw C' and D' equal and
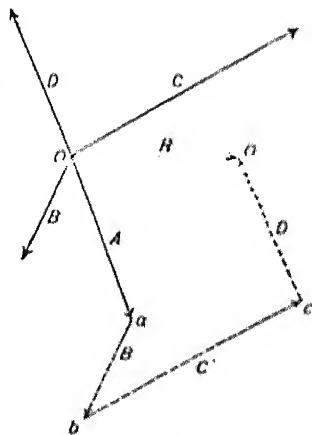
Fig. 2

parallel, respectively, to C and D, and it is clear that the result of the four displacements on a body originally at O would be to transfer it to O'. Therefore the *resultant* of the four displacements is the single displacement R, and this is so whether the component displacements occur simultaneously or successively. The particular order in which the several components are taken is quite immaterial.

This construction by which the resultant is found is called the diagram of displacements, it is perfectly general and applies whether the components are in the same plane or not.

20. Composition of Vectors. The above construction is a particular instance of the addition or composition of vectors. By a precisely similar process the resultant of any set of vectors may be obtained whether they represent forces, velocities, momenta, or any other quantities having direction as well as magnitude.

21. Resolution of Displacements. There is only one resultant displacement that can be found when the components are given, in whatever order they
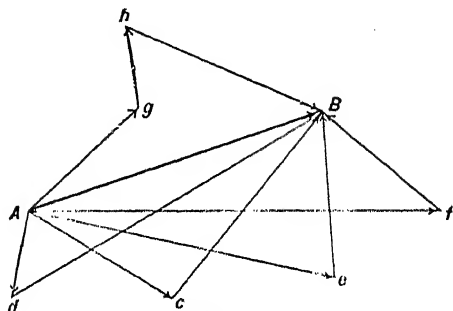
Fig. 3

may be taken. If it is required, however, to *resolve* a given displacement into its components, there are an infinite number of ways in which it may be done. For example, the displacement AB (Fig. 3) may be regarded as having Ac and cB as its components, or Ad and dB or Ae and eB, or it may be considered the resultant of the three displacements Ag, gh, hB. Or if any broken line whatever be taken starting at A and terminating at B, AB evidently will be the resultant of the displacements which are represented in amount and direction by the several parts of the broken line.

22. Resolving of Vectors. What has just been said of the resolving of a displacement into components is equally true of the resolving of any other vectors whatever into component vectors, and applies to the resolution of velocities, forces, etc.

23. Velocity. The velocity of a body is the rate at which it passes over distance in time. It is a vector quantity, its direction being as important as its amount. The term *speed* is familiarly used to express the *amount* of velocity without reference to its direction. Two bodies may be moving with the same speed, but if they are not going in the same direction their velocities are different.

This is the strict use of the word velocity; it is often somewhat loosely used to express merely the speed of motion.

24. Constant Velocity. When a body moves in a straight line always passing over equal distances in equal times it is said to have constant or uniform velocity. It is evident that the motion must be in a straight line, otherwise the *direction* of the velocity would not be constant.

In this case of motion if the length of any part of the path be divided by the time taken for the body to traverse that portion, the result is what is called the rate of *motion*, or the distance passed over per unit time, and is the same whatever part of the path may be chosen. It is this quantity which is the *speed* or the amount of the velocity.

Thus when a train is moving with constant velocity, the number of miles run in a given time divided by that time expressed in hours, is the speed in miles per hour.

25. Variable Velocity. When either the rate or direction of motion of a particle is changing, it is said to be moving with variable velocity. Thus the velocity is varying in case of a falling body which constantly gains in speed or in case of a railway train rounding a curve where the direction of motion is changing.

To understand what is meant by the *speed* of motion at a particular point when the velocity is constantly changing we

may consider a short portion *bc* (Fig. 4) of the path of the body having at its middle the point *a*

Fig. 4

at which the speed is to be determined. Divide the length of *bc* by the time taken by the body in traversing it. The result will be what may be called the *average* speed over that part of the path. If, now, the part chosen is taken smaller and
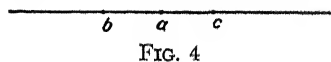
smaller, always having the given point at its center, the average velocities thus found will approximate more and more nearly to the true velocity at the given point, and that value which these successive approximations continually approach as a *limit*, as the distance *bc* approaches zero, is the speed of motion at the point *a*. At each instant a body has a certain speed, but it may not be constant even for the shortest interval of time that can be conceived.

So also with regard to the *direction* of motion. If the body moves in a curved path, its direction of motion at any point is the direction of the tangent to the curve at that point, and as the direction of the tangent constantly changes as we pass along the curve, so the direction of the velocity in such a case may be different at one point from what it is at a neighboring one, however near together the two points may be.

26. Composition of Velocities. If a body has at any instant several component velocities, the resultant velocity may be found by the vector diagram as in the case of the composition of displacements.
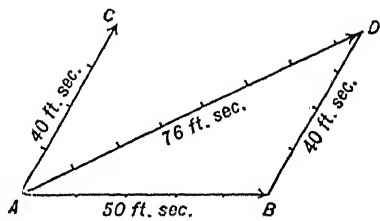
Fig. 5

For instance, suppose a ball is thrown in a moving railway car, it is required to find the velocity of the ball with reference to the earth. Let *AB* represent the velocity of the railway car, say 50 ft. per second, and let *AC* be the velocity of the ball as thrown obliquely across the car with a velocity of, say, 40 ft. per second. Then, laying off the vectors *AB* and *BD* with the proper relative direction and length, the resultant velocity is represented by the vector *AD*, which is found by measurement (using the same scale as in laying off *AB* and *BD*) to be 76 ft. per second, and this is the resultant speed of the ball relative to the earth. If the angle between *AB* and *AC* is given, the side *AD* of the triangle *ABD* may be calculated by trigonometry, using the formula

$$AD^2 = AB^2 + AC^2 + 2AB \cdot AC \cdot \cos CAB,$$

in which $\cos CAB$ replaces $- \cos ABD$.

27. Resolution of Velocities. Any given velocity may also be resolved into component velocities. For instance, suppose a man is rowing a boat with a velocity of 10 ft. per second in a
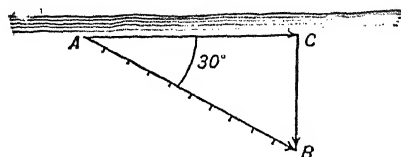
FIG. 6

direction making an angle of 30° with the straight shore of a lake, and it is required to determine how fast he is moving along the shore and how fast he is moving out into the lake. Let *AB* (Fig. 6) represent a velocity of 10 ft. per second. Draw *AC* parallel with the shore and making an angle of 30° with *AB*. Draw *CB* perpendicular to *AC*. Then *AC* and *CB* will represent two velocities, one parallel to the shore and one at right angles to it, whose resultant is *AB*. Therefore the boat may be regarded as having a velocity *AC* parallel with the shore and a velocity *CB* at right angles to it, and the amounts of these may be found by measurement, using the same scale as in laying off *AB*. Or we may calculate them by trigonometry, for

$$AC = AB \cdot \cos 30° \quad \therefore \quad AC = 8.66 \text{ ft. per second.}$$
$$CB = AB \cdot \sin 30° \quad CB = 5 \quad \text{ft. per second.}$$

It is frequently necessary in practice to resolve a velocity or other vector into two components which are mutually at right angles as in the case just discussed, and so this case, while one of the simplest, is one of much importance.

28. Acceleration. When the velocity of a body changes either in amount or direction the motion is said to be *accelerated*, and the change in velocity per unit time, or the time rate of change of the velocity, is called the *rate of acceleration* or simply the *acceleration*.

Change in velocity may always be thought of as due to the body receiving an additional component velocity which is compounded with the original velocity, the velocity after the change being the resultant of the two.

For example, in the upper diagram of figure 7 a body moving in a straight line is represented as having a velocity of v_1 at *A* and a greater velocity of v_2 at *B*. The gain in velocity is repre-

sented by the vector C which must be added to v_1 to give v_2. The average rate of acceleration between A and B is therefore found by dividing C, the increase in velocity, by the time taken by the body in passing from A to B. In the second diagram v_2 is less then v_1, and so the change in velocity is represented by the arrow D and is negative or opposite to the original motion. In the third case figured, the motion is along a curve and the velocity at B is not in the same direction as the velocity at A, but a velocity represented by E if compounded with v_1 will give v_2 as the resultant. The velocity E is, therefore, the change in velocity between A and B, and dividing it by the time during
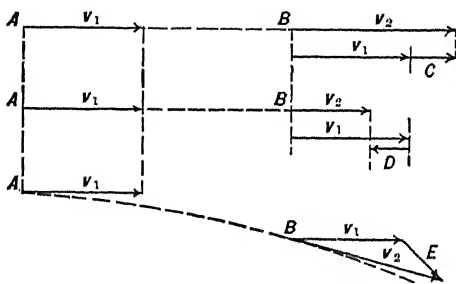
FIG. 7

which the change has taken place or the time of motion from A to B, the *average* rate of acceleration between A and B is found.

29. Acceleration in Rectilinear Motion. If the motion is in a straight line the velocity changes only in amount and not in direction, and the acceleration is calculated by dividing the change in speed during a given interval of time by the time interval. Or, expressing it in a formula,

$$u$$

where u represents the velocity at the beginning of the interval of time t while v represents the velocity at its end. This formula gives, in general, the *average rate of acceleration during the interval of time t*, but if the acceleration is constant it gives the actual rate.

Thus if a ball with a velocity of 50 ft. per second has, after

one-half second, a velocity of 40 ft. per second in the same direction, the average rate of the acceleration during the interval is

$$\frac{40 \text{ ft. per sec.} - 50 \text{ ft. per sec.}}{\frac{1}{2} \text{ sec.}} = -20 \text{ ft. per sec. per sec.,}$$

the negative sign indicating that the acceleration is opposite in direction to the original velocity and therefore the velocity is decreasing. The above is written in abbreviated form as follows:

$$\frac{40 \frac{\text{ft.}}{\text{sec.}} - 50 \frac{\text{ft.}}{\text{sec.}}}{\frac{1}{2} \text{ sec.}} = -20 \frac{\text{ft.}}{\text{sec.}^2}$$

30. Composition and Resolution of Accelerations. When a moving body has several different accelerations, as when a man in a railway car starts to walk in the car while the speed of the train is changing or while it is rounding a curve, the several accelerations may be compounded and their resultant found just as with other vectors. So also an acceleration may be resolved into two or more components.

PROBLEMS

1. A man walks $\frac{1}{2}$ mile in 10 minutes. What is his average velocity in feet per second?

2. A train has a velocity of 30 miles per hour; what is its velocity in feet per second?

3. A bicycle rider is traveling north at the rate of 10 miles per hour. If the wind is blowing from the east at the rate of 6 miles per hour, what is its apparent direction and velocity to the rider? Show the direction by a diagram.

4. A man rows a boat at the rate of 4 miles per hour, making an angle of 30° with the straight shore of the lake. How fast is he moving away from the shore?

5. Draw a diagram to scale showing the direction in which a man must row across a river in order to reach a point directly opposite, if he rows 3 miles per hour while the speed of the current is 2 miles per hour.

6. If the river in the last problem is $\frac{1}{4}$ mile broad, how long does it take to cross it as described, and what is the velocity of the boat relative to the shore?

7. A ball rolling down an incline has a velocity of 60 cm. per sec. at a certain instant, and 11 seconds later it has attained a velocity of 181 cm. per sec. Find its acceleration.

8. A body having an initial velocity of 60 ft. per sec. has an acceleration — 32 ft. per sec. per sec. Find its velocity at the end of 1, 2, and 3 seconds.

9. A railroad train having a velocity of 40 miles per hour is brought to rest in 1 minute. Find the acceleration in feet per second per second.

First Principles of Dynamics

31. First Law of Motion. When a ball on a table starts to roll, experience convinces us either that the table is not level or that some external force other than gravitation has caused the motion. On the other hand, when we see a ball that has been set rolling on a level table, gradually losing its speed, we are equally satisfied that there is some force resisting its motion. For it is found in such a case that if the table is made smoother and if air resistance is gotten rid of, the ball loses speed much more slowly than before. We are thus satisfied that if there were no force resisting its motion the speed of the ball would remain unchanged.

This conviction, arrived at through experience, was clearly enunciated by Sir Isaac Newton in the first of his celebrated Laws of Motion, published in his *Principia*, in 1686.

First Law of Motion : Every body continues in its state of rest, or of moving with constant velocity in a straight line, unless acted upon by some external force.

32. Discussion of the First Law of Motion. The first law asserts that force is not required to *keep* a body in motion, but simply to *change* its state of motion. After a railroad train has attained a constant speed the entire force of the locomotive is spent in overcoming the various resistances that oppose the motion, such as friction of wheels and bearings and air resistance. But for these the train would maintain its speed without aid from the locomotive.

Therefore, when any object is observed to be at rest or moving with constant speed in a straight line, we conclude either that no external force acts upon the body, or that whatever forces act are so related as to neutralize or balance each other.

Since we measure equal times by the equal angles through which the earth has moved, the law that freely moving bodies move through equal distances in equal times may seem simply a consequence of the mode of defining equal times and without any physical significance. But the statement of the law really asserts the physical fact that *in case of any two bodies whatever, unacted on by external forces, while one body moves through successive equal distances, the distances traversed simultaneously by the other body are also equal among themselves.*

That this is true whatever the nature of the bodies concerned is a fact of nature that rests on experience, and cannot be regarded as known *a priori*.

33. Inertia. The property, common to all kinds of matter, that no material body can have its state of rest or motion changed without the action of some force, is known as *inertia*. The amount of force required to produce a given change in the motion

of a body depends both on the body and on the suddenness of the change to be produced.

Any force however small can give as great a velocity as may be desired to any mass however great, provided it acts for a long enough time.

If a weight rests on a sheet of paper on a table it may be drawn along by means of the paper, for there is friction between the two and it requires a certain force to slip the one over the other. *If, therefore, we do not attempt to accelerate the weight too rapidly,* the friction will move the weight along with the paper. But if we attempt to start the weight suddenly, or change its velocity suddenly while moving, the paper will at once slip from under it, for the force required to produce the sudden change of motion is greater than the friction between the two.

FIG. 8.   Inertia

Again if a 10-lb. weight (Fig. 8) is hung by a cord from a fixed support and if it is drawn steadily downward by a piece of the same cord attached to it underneath, the cord will break *above* the weight, for the force exerted by the lower cord upon the weight will cause it to move downward, straining the upper cord with the combined force due both to the weight and the pull. But a sudden pull will break the cord *below* the weight. For in consequence of its inertia the weight cannot be set in motion as suddenly as the cord is pulled without the
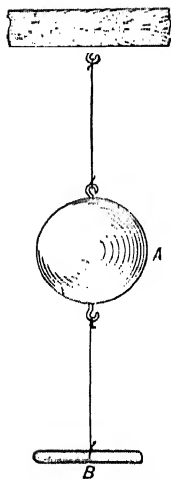
exertion of a greater force than the cord is able to bear, so that the cord breaks even before the weight has moved downward enough to strain the upper cord to the breaking point.

The complete statement of the effect of a force upon the motion of a body is embodied in Newton's Second Law of Motion, and will be discussed when we take up the study of Kinetics (§ 93).

34. Measure of Mass. The inertias of bodies may be compared quantitatively by the amounts of force required to accelerate them at the same rate, and when they are thus compared it is found that *the inertia of a given portion of matter is always the same and cannot be increased or diminished by any known process.* Consequently the inertia of a body is the sum of the inertias of its several parts, the inertia of two quarts of water is twice that of one quart whether they are combined or separate.

It is for this reason that the quantity of matter in a body, or its *mass* as it is called, is measured by its inertia as compared with that of some standard piece of matter taken as the unit of mass.

Therefore, *masses are said to be equal which acquire equal velocities when acted on by equal forces for the same length of time.* For example, suppose two masses are drawn side by side over a frictionless surface by two spring balances at such a rate that each balance is kept constantly stretched, say to the 4-oz. point, so that they exert equal forces; then, if the masses after starting together keep pace with each other, they are acquiring velocity at the same rate and consequently are equal.

Of course such an experiment serves chiefly to illustrate what is meant by saying that equal masses have equal inertias, for it would be impossible to directly compare masses in this way with any degree of accuracy.

The actual comparison of masses is accomplished with great accuracy by *weighing;* for it is found that masses which have equal inertias have also equal weights, provided they are weighed in a vacuum at the same point on the earth (§ 102).

35. Measure of Force. A force may be measured in three ways:

1. *By the weight that it can support.* This is the gravitation method.

2. *By its power to strain an elastic body,* as in the ordinary spring balance.

3. *By its power to give motion to a mass.* This is the dynamical method.

The first method is very convenient and forms the basis of most measurements of force in engineering and ordinary life, but it has the disadvantage that the force required to support a pound weight varies from place to place on the earth.

The second method is convenient for comparing forces, but the elastic properties of one substance differ from those of another, besides being dependent on temperature and physical condition, so that a standard force could not be preserved or accurately defined by this method.

The third method is very important because it is used to define the standard of force in physics. It is difficult to apply except indirectly, but furnishes a unit of force which depends only on the inertia of matter and is, therefore, absolutely invariable and well suited to be a standard force.
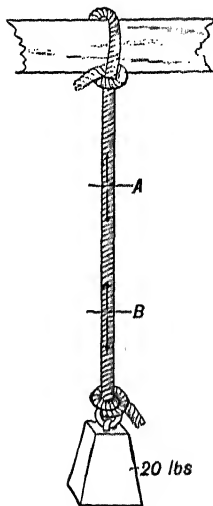
36. Equal Forces. Two forces are said to be *equal* when the velocity of a given mass is increased at the same rate per second by one force as by the other. When two such forces act in opposite directions on a given mass they neutralize or balance each other so far as any effect on the *motion* of the mass is concerned. Thus when a cord is stretched horizontally between two springs, the forces exerted by the springs are equal and opposite so long as the cord remains at rest or moves with uniform velocity.

37. Stress. When a weight is supported by a uniform cord, every part of the cord is stretched, and if the weight of the cord itself is so small that it may be neglected, the stretch of every inch of it is the same whether it is near the upper or lower end and whatever may be the total length. The section *AB* (Fig. 9) is pulled up by the cord above *A* and is pulled down by the cord below *B* and is.

FIG. 9

therefore, stretched until the contractile force of its own elasticity balances the external stretching force. In this way every portion of the cord is subject from without to an external stretching force, and it exerts in opposition to this an internal contractile force and is said to be *in a state of stress*. In this case the stress is called *tension*, and every· portion is subject to forces which tend to elongate it. When a weight is supported on a vertical rod or column the whole support is in another state of stress called *pressure* or *compression*, for the forces that act on any part of the rod tend to shorten it. Besides *tension* and *pressure* there is a third kind of stress, called *shearing stress*, which tends to distort or force out of shape the parts of a body. This is the stress in a rod that is being twisted. But the further discussion of this matter must be left until the elasticity of bodies is considered (§ 242).

38. Action and Reaction. Every stress has a double aspect. Thus when a weight rests on a table the force between the two may be regarded as a pressure down on the table or an upward push against the weight. When a cord is supporting a weight, at every cross section in the cord there is a downward pull on the cord above the section, and an upward pull on the cord below and these two are exactly equal. When a magnet attracts a piece of iron the force may be regarded as drawing the iron toward the magnet or the magnet toward the iron.

These two aspects of a stress are known as the action and reaction; they are exactly equal and opposite. This fundamental fact was stated by Newton as the Third Law of Motion.

Third Law of Motion: To every action there is an equal and opposite reaction.

39. Discussion of Third Law of Motion. When a weight rests upon a table it is pushed up with a force equal to that which it exerts upon the table. The table, therefore, presses a heavy weight upward with more force than it exerts on a small weight. The only limit of the power of the table to react is its strength. It is instructive to consider what happens when a weight heavy enough to crush the table is placed upon it. As it is lowered upon the table it presses more and more until the limit of the table's power of resistance is reached, when in breaking down it begins to move away from the weight at such a rate that the reaction

which it exerts is at every instant exactly equal to the pressure to which it is subjected by the weight. For if a body moves away fast enough from another which is pressing upon it the pressure may be diminished to any extent.

When a ball is struck by a bat the force upon the bat at every instant while they are in contact is the same as that which the bat exerts upon the ball.

40. Composition and Resolution of Forces. When several forces act simultaneously on a particle *the single resultant force may be found by the diagram of vectors* (§ 19) just as in case of accelerations. This follows from the fact that when a particle is acted upon by a force it is accelerated in the same direction in which the force acts and by an amount which is proportional to the force. (See §§ 93-94.) If several forces act upon the particle simultaneously, each force produces a component of acceleration proportional to it. Thus to every acceleration component a corresponding force component exists, and the resultant force is found by a diagram of vectors representing forces, just as the resultant acceleration is found by a diagram of vectors representing acceleration components. Since the accelerations are proportional to the forces the vector diagrams are similar and the resultant force is in the same direction as the resultant acceleration produced by it.

Thus any force may be considered as the resultant of two or more component forces, and these components may be found just as the components of an acceleration are found.

41. A Special Case. Suppose the vector *AB*, five units long, represents a force of five pounds, acting obliquely on a block of wood resting on a table, and it is required to find how much force is pressing the block against the table and how much is urging it along its surface. The vector *AB* may be resolved as just explained into the two components *AC* and *CB*, where *AC* represents the force pressing the block against the surface and *CB* represents the force pushing it along the surface. The amount of these components may be determined either by direct measure-

Fig. 10. Oblique force on block

ment, using the same scale as in laying off AB, or it may be calculated as follows: If F is the amount of the force AB and if a is the angle between AB and the table top, then

$$AC = F \cdot \sin\ a.$$
$$CB = F \cdot \cos\ a.$$

II. STATICS

EQUILIBRIUM OF A PARTICLE

42. Equilibrium. Before taking up the study of the motions of bodies as determined by forces, we shall consider some cases in which the various forces concerned are so related as to balance each other, so far as the motion of the body on which they act is concerned.

A body is said to be in equilibrium when any forces which act on it are so related that the body is not accelerated. Thus a body at rest is in equilibrium, also a body moving with constant velocity in a straight line, also a body turning with constant speed of rotation about an axis through its center of mass, as in case of a well-balanced wheel, also when such a wheel is not only turning with constant speed, but moving along with constant speed. Thus a wheel rolling in a straight line along a level surface or a wheeled vehicle like a car on a straight level track is in equilibrium if moving with constant speed.*

We shall first consider the equilibrium of a *particle* or a body so small that the forces acting on it may all be considered as acting at one point. Afterward the conditions of equilibrium of an extended rigid body will be taken up.

Whether a body is to be treated as a particle or not depends on circumstances. For instance, in astronomy the sun and planets are treated as particles when their shapes and distribution of mass do not affect the question considered.

43. Equilibrium of a Particle. A particle is in equilibrium when the resultant of the forces acting on it is zero. Evidently

---

* When moving as just described, a wheel *considered as a whole* is in equilibrium, but its particles are *not* in equilibrium, for they move in circles and are therefore accelerated (§ 28).

in this case the diagram of the forces must be a closed triangle or polygon.

For let the forces acting in a given case be *abc* (Fig. 11), then if we draw the diagram of forces as in the lower part of the figure and if the vectors *abc* form a closed triangle, as shown, the resultant is zero and the particle is in equilibrium.

So also in case of any number of forces in equilibrium, *the diagram of forces, formed by drawing successively the several vectors representing the forces, must be a closed polygon;* that is, the last vector drawn must terminate at the starting point.

An example of four forces in equilibrium is shown in figure 12, the several forces *abcd* forming a closed polygon.

It is interesting also to observe that if we resolve each of these forces into two components, one directed toward the top or bottom of the page and the other sidewise, as, for example, *b* is resolved into *b'* and *b''*, *c* into *c'* and *c''*, etc., we find that the components *a'b'* directed from left to right exactly balance the components *c'd'* directed from right to left, so also the upward components *a''* and *d''* are together equal to the sum of the downward components *b''* and *c''*.

Fig. 11. Three forces in equilibrium

In the above diagram for convenience all the four forces have been represented in one plane. This restriction is not necessary, the same construction is the test of equilibrium in whatever directions the four forces may act.

44. Illustrations. If three cords joined at *P* suspend weights of 3, 4, and 5 lbs. respectively, those supporting the 3 lb. and 4 lb. weights passing over frictionless pulleys as shown in figure 13, then the point of junction *P* will assume a definite position to which it will return if pushed aside and the cords *PA* and *PB* will be at right angles to each other.

For the point *P* is in equilibrium under the three forces 3, 4, 5, and therefore the force diagram must be the closed triangle *PCD*, the three sides of which are in the ratio of 3 : 4 : 5. But such a triangle is right-angled and therefore the force 4 and force 3 must be at right angles to each other.

Suppose a cord is fastened at *A* and *B* and is then stretched by a weight

W hung at P (Fig. 14). As before, the diagram of forces is PCD, where PC represents the stress on the cord between P and B while CD represents the stress on AP, and DP represents the weight W. Evidently the more nearly AP and PB are to being in a straight line the larger will CD and PC be in comparison with the force W which is represented by DP. So that a
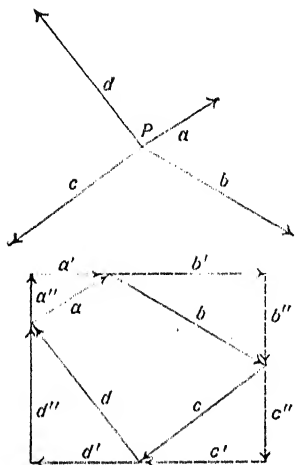
FIG. 12. Four forces in equilibrium   FIG. 13
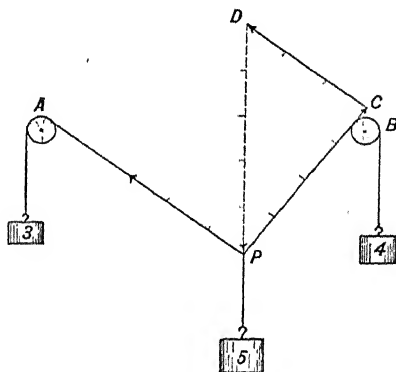
comparatively small pull down at P, if APB is nearly straight, may produce a force great enough to break the cord between A and B. Thus the stress brought to bear on hammock ropes may be much greater than the weight of the person supported if it is hung with insufficient sag.
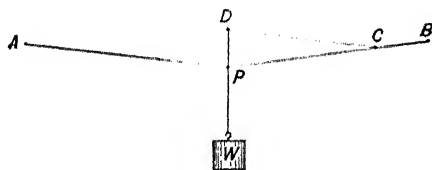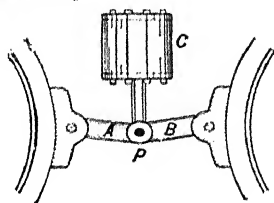
FIG. 14   FIG. 15

The student may easily determine under what conditions the stresses on AP and PB will be each equal to the weight W.

The jointed device used in hand printing presses, and shown in figure 15 as applied to the brakes between the drivers of an old-fashioned locomotive, illustrates the same principle. Here when compressed air is admitted to the cylinder C the piston is forced upward, thus straightening the two connecting pieces A and B, thereby forcing the two brake shoes against

the wheels with a force which is greater the more nearly the connecting pieces are pulled into a straight line.

45. Bridge Stresses. Let it be required to find the stresses, tensions or pressures, in case of the various parts of the truss which is shown in figure 16, supporting at its center the weight H.

Consider what forces are acting on the end of the truss at A. It will be shown later (§ 54) that in such a case half the total weight will be borne by one abutment and half by the other. The end of the truss at A therefore presses down on the abutment with a force equal to $\frac{H}{2}$, if we neglect the weight of the truss itself. Now A is in equilibrium under the three forces represented by the arrows, P indicating the upward pressure of the abutment, S representing the oblique downward thrust of the strut AB, while T represents the inward pull of the tie rod AC; therefore the diagram of these
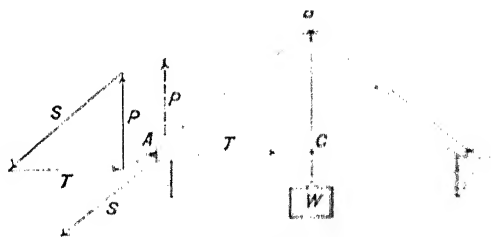
FIG. 16. Bridge truss

forces must be a triangle as shown above. But this triangle is similar to the triangle ACB, for the sides are respectively parallel, and so the forces, P, S, and T are in the same proportion as the sides BC, AB, AC, and since the pressure P is equal to $\frac{W}{2}$, the other forces are at once known by proportion when the sides of the triangle ABC are known.

In the somewhat more complicated case shown in figure 17 where the total weight of 10 tons is supported between the two abutments the upward pressure P will equal 5 tons. The stresses on AB and AC may be found as in the preceding case, but to find the stress on the rod BC we must make a diagram of the forces under which the point B is in equilibrium as shown in the diagram.

46. Crane Problem. A weight of 100 lbs. is suspended from a crane of dimensions shown in the figure. It is required to find the tension on the tie rod AB and the compression on the strut BC.

The point B is in equilibrium under three forces, the downward weight $W = 100$ lbs., the pressure P of the strut which acts outward, and the tension T of the tie rod which acts in the direction BA. The diagram of forces must therefore be a triangle with sides parallel to the directions of the three forces as shown in the figure.

NOTE: In the illustrative problems of Figs. 15–18 the diagram of forces can be made use of as shown, only when the compressional forces lie exactly along the compression members. This is true if the weight of the compression members is neglected, and if they are hinged at each end so they cannot transmit bending forces. In the problems of this type which follow, these conditions are assumed to exist. In the construction of large bridges, however, the weight of the individual members may be an important consideration.
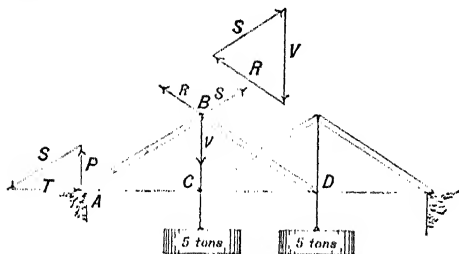
FIG. 17
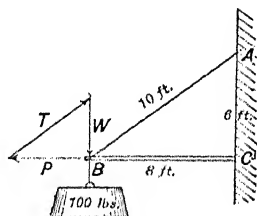
FIG. 18.   Crane

PROBLEMS

1. A force of 300 grams and a force of 400 grams act at right angles to each other on the same point. Find the single force to which they are equivalent and also its direction, by a diagram.

2. Four forces of 3, 4, 5, and 6 lbs. act on the same point in directions east, northeast, north, and northwest, respectively. Construct the force diagram and find by measurement the amount of the resultant and the angle which it makes with the north line.

3. Three cords fastened together at a point free to move, have tensions 60, 70, and 80 grams, respectively. Construct the force diagram and find by measurement the angles between the cords when at rest.

4. Two forces of 10 lbs. each act upon a single point in such a way that they are equivalent to a single force of 10 lbs. Find the angle between their lines of action.

5. How much force must be exerted at an angle of 45° to the top of a table to push along a weight when the frictional resistance to be overcome is a force of 2 kgms.?

6. When a force of 20 lbs. is required to draw along a sled by a rope making an angle of 30° with the ground, find the force moving the sled forward, and the force diminishing its pressure on the ground.

7. A weight of 2 lbs. hung from a nail by a cord 30 in. long is pushed aside by a horizontal force of 1 lb. How far will it be moved away from the vertical line through its point of support?

8. A 32-gram weight is hung by a cord 60 cm. long from a point on a vertical wall. How far will it be pushed out from the wall by a force of 24 grams acting perpendicular to the wall?

9. Make a diagram showing the angle between the two ropes of a hammock when the tension on each rope is twice the weight of the person in the hammock.

10. A rope supporting a weight of 180 lbs. at its middle point is hung between two hooks which are on the same level and 18 ft. apart. If the middle point sags 3 ft. below the level of the hooks, find the force on each hook.

11. In case of a crane, like figure 18, in which the horizontal strut is 5 ft. long and the vertical distance AC is 3 ft., find the tension on the oblique tie rod and the pressure on the strut when a weight of 270 lbs. is supported.

12. Suppose the wall in figure 18 overhangs so that A is 6 ft. vertically above a point 1 ft. to the left of C on the bar BC, which is horizontal and 9 ft. long. Find the stresses on AB and BC when a weight of 240 lbs. is suspended at B.

EQUILIBRIUM OF A RIGID BODY

47. Equilibrium of a Rigid Body. A rigid body is in equilibrium when its velocity of translation is not changing in any direction and when its velocity of rotation is not changing about any axis.

Or, in other words, *a rigid body is in equilibrium when it has no acceleration either of translation or rotation.*

48. Condition for Translational Equilibrium. In order that there may be no *translational* acceleration, the relation between the forces acting on the body must be exactly the same as that required for the equilibrium of a particle. For if there is to be no acceleration in any direction the resultant force in any one direction must be zero, and this is evidently the case *when the diagram of forces is a closed polygon.*

49. Case of Two Forces. The relation which must hold between the forces in order that there may be no *rotational* acceleration may be most easily reached through the study of some simple cases of equilibrium. That a body may be in equilibrium under two forces it is necessary that the two forces P and Q (Fig. 19) should be equal and opposite in order to satisfy the condition of no translational acceleration as just shown. In order that there may be no tendency of the forces to rotate the

body it is clearly necessary that they shall act *in the same straight line*, as shown in the figure. The only effect of the forces applied at A and B in the figure is to compress the body between these points.

50. Resultant of Two Oblique Forces in the Same Plane. Let P and Q represent two forces acting at A and B upon an extended body, and let their lines of action when produced intersect at C. A force equal and opposite to P if applied at C will exactly balance P, as shown in the preceding paragraph, and a force equal and opposite to Q, also applied at C, will balance Q; therefore a single force R equal and opposite to the resultant of P and Q as found by the triangle of forces, will, if applied at
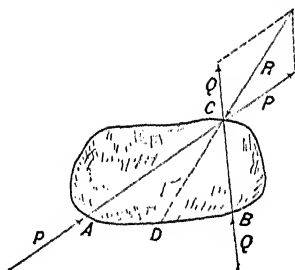
FIG. 19   FIG. 20

C, exactly balance both P and Q and produce equilibrium. Evidently the force R may be applied to the body at any point in its line of action CD. The resultant of P and Q is, therefore, a force equal and opposite to R and acting along the line CD.

51. Moment of Force. In the case of equilibrium just discussed we may imagine the force R to be produced by pressure against a *pivot* fixed somewhere in the line CD, say at E (Fig. 21).

The forces P and Q then balance each other, so far as causing rotation about the axis at E is concerned. Two forces so related to any axis are said to have equal and opposite *moments* with respect to that axis.

Common experience shows us that the farther the line of action of a force is from the pivot or axis, the greater will be its ability to rotate the body about that axis. Thus in opening a heavy gate we take hold of it as far as possible from its hinges and pull at right angles to the gate. A pull in line with the hinges would have

no effect to turn it whatever. The moment of a force to turn a body about an axis depends, therefore, both on the amount of the force and the distance of its line of action from the axis.

The moment of a force with reference to a given axis is a measure of its ability to produce rotation about that axis, and is numerically equal to the product of the force by the perpendicular distance from the axis to the line of action of the force.
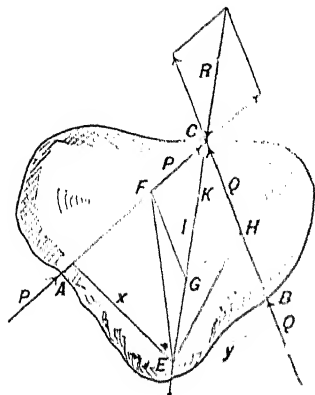
The following proof of this statement is given for the special case of two forces lying in the same plane.

Let x be the perpendicular distance from E to the line of action of the force P, and let y be the perpendicular distance from E to the line of action of Q. We shall now show that in this case, where the moments of P and Q about the axis E balance each other, $Px = Qy$.

Construct the parallelogram $CFGH$ having $CF = P$ and $CH = Q$. Draw FE and HE; then the area of the triangle CFE is equal to $\frac{1}{2}Px$ for P is its base and x is its altitude. So also the area CHE is equal to $\frac{1}{2}Qy$. But the two triangles CFE and CHE have equal areas, for they have a common base CE and equal altitudes HI and FK; therefore, $Px = Qy$.

FIG. 21

It has thus been shown that when the axis E lies on the line CG the two forces P and Q have equal and opposite moments about it and also in that case $Px = Qy$. *These products Px and Qy may, therefore, be taken as representing the abilities of the forces to produce rotation about the given axis*, and are, therefore, used to measure the moments of the forces.

52. Second Condition of Equilibrium. The second condition of equilibrium for an extended rigid body is that the various forces must be so related that there is no rotational acceleration about *any* axis in the body.

Since the ability of a force to produce rotation is measured by its moment, *this condition is satisfied when the sum of the moments of the forces tending to produce clockwise rotation about any axis*

whatever is equal to the sum of the counter-clockwise moments about that axis.

53. Forces in One Plane.   When the forces acting on a body all lie in one plane, such as the plane of the paper in figure 22, they can have no tendency to rotate the body except about an axis at right angles to that plane.

In this case if the diagram of forces is a closed polygon showing that there is no translational acceleration, and if the clockwise and counter-clockwise moments are equal about some one axis at right angles to the plane of the forces, then the body is in equilibrium and the resultant moment of the forces is also zero about any other axis that may be chosen in the body.
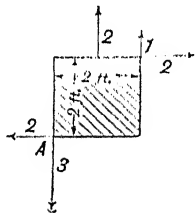
Fig. 22

For example, in case of a board 2 ft. square, with forces applied to it as shown in figure 22, the diagram of forces is a closed figure, as the student may easily verify. The forces, therefore, balance so far as *translation* is concerned.

Now calculate the moments of the forces about an axis perpendicular to the plane of the forces, say through the point *A*. Designating clockwise moments *plus* and counter-clockwise *minus*, and taking the forces in order beginning at the top, we have the moments

$$
\begin{aligned}
- 2 \times 1 &= -2 \\
- 1 \times 2 &= -2 \\
+ 2 \times 2 &= +4 \\
3 \times 0 &= 0 \\
2 \times 0 &= 0
\end{aligned}
\Biggr\} \text{ Sum of the moments } = 0.
$$

Therefore the board is in equilibrium under these forces and consequently the sum of their moments will be zero if reckoned for any axis whatever.

Compute in this way the moments about an axis through the center of the board and show that their sum is zero.

54. Three Parallel Forces.   When a bar is in equilibrium under three parallel forces, as in figure 23, to satisfy the condition of no translational acceleration the *up* forces must be equal to the *down* forces, or $P + Q = R$. While to satisfy the second condition, that the moments of the forces shall balance, we have $Px = Qy$, for these are the moments about the point *B*, and *R* has zero moment about that point. Or we may take moments about *A* and find $Rx = Q(x + y)$. If moments are

taken about any point other than A, B, or C, there will be three moments to reckon. If, for example, the point D is taken as the axis the clockwise moment of Q must be equal to the sum of the counter-clockwise moments of P and R.

55. Parallel Forces in General. Any case of equilibrium with parallel forces may be discussed in a similar way, two conditions being met, namely, the sum of the forces acting in any one direction must be equal to the sum of the forces in the opposite direction, and the sum of the clockwise moments about
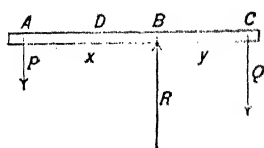
Fig. 23   Fig. 24

any axis must be equal to the sum of the counter-clockwise moments.

56. Illustration. A certain bar having no weight is acted on by four forces as shown in figure 24, forces of 4 lbs. and 2 lbs. acting upward and 3 lbs. and 5 lbs. acting downward, and it is required to find the single force necessary to produce equilibrium and the point on the bar where it must be applied. Since the total upward force is 6 while the downward force is 8, the required force F must be an upward force 2 to satisfy the first condition of no translational acceleration.

This force must be applied at such a point on the bar as to satisfy the *second condition*, and make the clockwise moments balance the counter-clockwise moments about any axis. Take an axis through P, for instance. The moments about P are

$$\begin{array}{rcl} 3 \times 0 &=& 0 \\ -2 \times 1 &=& -2 \\ +5 \times 2 &=& +10 \\ -4 \times 3 &=& -12 \end{array} \qquad \text{Sum} \qquad 4 \text{ counter-clockwise.}$$

Therefore, to produce equilibrium the applied force 2 must produce a clockwise moment 4. Since it must also act upward, it must be applied at a distance 2 to the *left* of P, and consequently the bar must be extended 2 feet in that direction.

Any point whatever on the bar might have been taken as the origin of moments, and the reader should show that the same conclusion is reached taking moments about some point such as C.

57. Couple and Torque. If in the case just treated the upward force 4 is changed to 6, we have a case that calls for special consideration. The upward forces are exactly equal to the downward forces and yet the bar is not in equilibrium, for taking moments about P we find that the clockwise moment is 10 while the counter-clockwise moment is $2 + 18 = 20$. Here, then, there is *a combination of forces that does not tend to produce translation, but simply rotation.* Such a combination is known as a *couple,* and its moment is commonly known as a *torque.* It cannot be balanced by any single force, for any force applied either upward or downward would cause translation. *A couple*

Fig. 25.

can be balanced only by another couple having an equal and opposite moment, or torque.
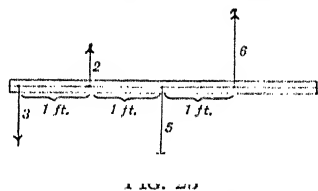
The simplest case of a couple is when two equal parallel forces act in opposite directions not in the same straight line. For instance, the forces FF, figure 26, constitute a clockwise couple the moment of which is Fx where x is the distance between the lines of action of the forces. *The moment of a couple about any axis is the same as about any parallel axis.* For, take an axis perpendicular to the paper and through P at a distance y from the nearer force, then the moments are Fy counter-clockwise and $F(x + y)$ clockwise, hence subtracting we have Fx clockwise, as the resultant of the two.
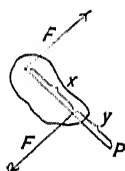
Fig. 26. Couple

The moment of such a couple about an axis perpendicular to the plane in which the two forces lie is, therefore, measured by the product of the amount of either force by the perpendicular distance between their lines of action.

To produce equilibrium, then, in the case under consideration a couple having a clockwise moment 10 must be applied to the bar, and it may be applied at any point we choose. The following figure illustrates different modes of producing equilibrium in this case.

In every case of equilibrium the forces acting may be resolved into a number of balancing couples.

58. Center of Gravity. The weight of a mass is the force with which it is drawn toward the earth. All parts of a body have weight and so the weights of the several parts into which a body may be conceived to be divided constitute a system of parallel forces acting downward toward the earth. The resultant of this system of forces is a single force equal to their sum and is the total weight of the body.
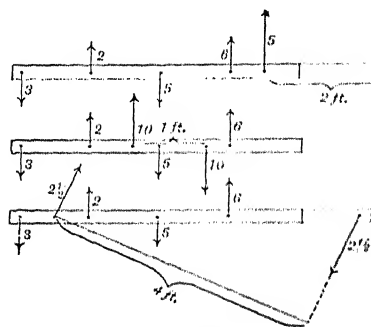
FIG. 27

It may be proved that there is a certain point in the body through which the resultant force due to weight always acts whatever may be the position of the body. This point is called the *center of gravity* of the body.

In all problems that involve the weight of a body we may ignore the fact that the weight is distributed throughout the body, and treat it as a single force applied at the center of gravity.

59. Proof of Center of Gravity. Let M and m be the masses of two parts of a body and let the line joining them be inclined as shown in figure 28. Since the weights of masses are proportional to the masses themselves (§ 38), the single upward force necessary to balance the weights of the two masses must be applied in the vertical line AB, so situated that $Mx = my$. But AB intersects at P the line joining the two masses, dividing it into the two segments a and b which, by similar triangles, are in the same ratio as x and y, and consequently $a : b :: m : M$;

FIG. 28

and since this ratio does not depend on the inclination of the line joining M and m, it follows that the balancing force must pass through the point P whatever the inclination may be. P is, therefore, the center of gravity of M and m. Now conceive the masses M and m concentrated at P and find similarly a point P' through which the resultant weight of $(M + m)$ and of another

mass m' must always pass.   Continue in this manner until account has been taken of all the masses into which we may conceive the body to have been subdivided.   The point through which the final resultant passes is the center of gravity of the body.

60. Center of Mass and of Inertia. The center of gravity as has just been explained is determined by the distribution of
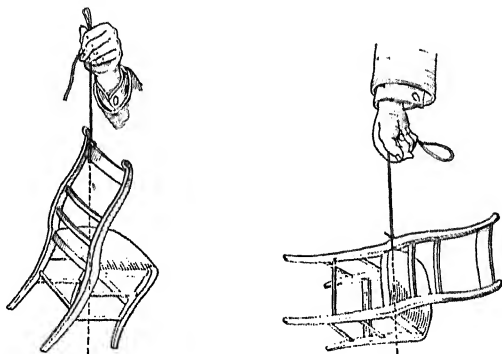
FIG. 29

mass in a body or system of bodies.   It has certain remarkable properties quite independent of weight, and therefore is also called the *center of mass* or *center of inertia* of the body or system.

For example, a freely rotating body like a spinning projectile will always rotate about an axis through its center of mass.

61. Position of Center of Gravity. *When a body is hung by a cord or balanced on a point the center of gravity must be in the vertical line passing through the point of support.* For two equilibrating forces must act in the same straight line. If, therefore,
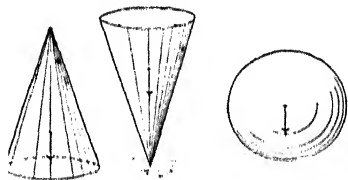
FIG. 30

a body is hung first from one point and then from another the intersection of the two lines thus determined marks the position of the center of gravity, as shown in the figure, where it is seen to be a point outside the actual substance of the chair (Fig. 29).

The center of gravity of a uniform bar is at its center; in a uniform thin plate, square, rectangular, or in the form of a

parallelogram, it is at the intersection of the diagonals. In case of any homogeneous symmetrical body it lies in the plane of symmetry. Thus it is at the center of a sphere, circle, or circular disc and at the center of a cube (Fig. 30).

62. Equilibrium under Gravity. That a body may be in equilibrium under gravity, it must be supported by a force equal
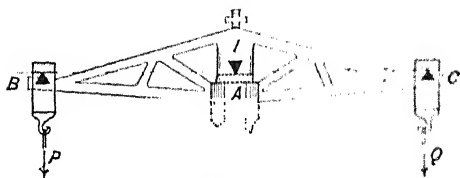
FIG. 31

to its own weight and acting upward through its center of gravity. Thus the two cones and sphere shown in the figure as resting on a level table are in equilibrium, the upward force being supplied by the reaction of the table. But the first cone is said to be in *stable* equilibrium, because if slightly tipped it will fall back to its original position. The second cone is said to be in *unstable* equilibrium because if disturbed it will fall away from its original position, while the sphere is said to be in *neutral* equilibrium because it remains in equilibrium when displaced. *It will be observed that in the first case the center of gravity of the cone is raised when it is tipped; in the second it is lowered, and in case of the sphere it is neither raised nor lowered.* The weight of a body being considered as acting at its center of gravity will always cause that point to move *down* or toward the earth unless opposed by some other force. In case of a loaded wagon on a hillside the vertical line through its center of gravity may remain between the wheels if the center of gravity is low, when if it were high the line of action of the weight might fall outside the wheel base causing the load to overturn.

63. Balances. The beam of a balance rests on a sharp steel "knife edge" A, while the pans are hung on the knife edges B and C. These three knife edges are rigidly fixed in the beam and should be parallel, and in the same plane, and the arm AC should be equal to the arm AB (Fig. 31).

Now if A, B, and C are all in the same straight line, and if the weight P is greater than the weight Q, the balance will tip entirely over unless some

counteracting force is available. This is found in the weight of the balance beam itself which is so adjusted that its center of gravity does not lie exactly on the edge A, but slightly below it, say a distance x. Then when P is greater than Q the balance beam inclines and its center of gravity is displaced until the restoring moment due to the weight of the beam W, acting down through its center of gravity just balances the deflecting moment due to the difference between P and Q. The deflection is therefore very nearly proportional to $P - Q$, and the greater the deflection for a given difference between P and Q, the greater is said to be the *sensitiveness* of the balance.

PROBLEMS

1. A wooden bar 5 ft. long and weighing 2 lbs., the ends of which are supported by spring balances, has a 10-lb. weight hung on it 2 ft. from one end. Find the force exerted on each balance.

2. A beam 20 ft. long is carried by three men, one at one end and the other two supporting it between them on a cross-bar at such a point that each man carries an equal weight. Find where the cross-bar must be placed.

3. A man standing on a uniform beam, 16 ft. long and weighing 120 lbs., at a point 1 ft. from its end causes it to just balance as it lies horizontally across a support 4 ft. from that end. What is the man's weight?

4. At what point on a pole must a weight of 52 lbs. be hung so that a boy at one end may carry $\frac{4}{9}$ as much as the man at the other end, and how much does each carry? Neglect the weight of the pole.

5. If the pole in problem 4 is uniform and weighs 10 lbs., where must a 50-lb. weight be hung so that the man may carry twice as much weight as the boy?

6. Find the center of gravity of a uniform bar weighing 6 lbs. and having a 2-lb. weight on one end and a 7-lb. weight on the other.

7. Forces 2 and 4 acting upward are applied to a horizontal bar at 2 ft. and 4 ft. from the left-hand end, respectively, also forces 3 and 1 acting downward are applied at 1 ft. and 5 ft. from the same end. Find amount and point of application of a single force producing equilibrium.

8. How produce equilibrium in problem 7 when an additional force 2 acts downward at a point 3 ft. from the left-hand end of the bar?

9. A board 2 ft. square is acted on by five forces applied at the same points as shown in figure 22, but the forces instead of being 2, 1, 2, 3, 2, beginning at the top, are 4, 2, 4, 5, 3, respectively. Find the direction and the amount of the force needed to produce equilibrium and how far from the center of the board its line of action must lie.

10. A ladder standing 6 ft. from a smooth vertical wall rests against it at a point 30 ft. from the ground. If the ladder weighs 60 lbs. and its center

of gravity is $\frac{1}{3}$ of its length from the bottom, find the force with which it presses against the wall, also the amount and direction of its force against the ground; that is, its vertical and horizontal components.

NOTE: The force between ladder and wall must be perpendicular to the latter if there is no friction between them.

11. When a man weighing 150 lbs. is halfway up the ladder in problem 10, find the pressure of the ladder against the wall and also the two components of its force against the ground.

12. When is the ladder in problem 11 more liable to slip, when a man is near the top or bottom, and why?

WORK AND ENERGY

64. Work. A man digging a ditch is said to work, so also a team of horses drawing a load, and a carpenter supporting temporarily the end of a beam may also, in ordinary speech, be said to be working; for in each case a useful end is secured by the exertion of force.

But there is a difference between these cases. In the first two there is motion and a permanent change is effected; while in the third case the beam is not moved but simply supported, and any prop would have served as well as the carpenter.

In physics the term *work* is restricted to such cases as the first two where motion results from the action of force, and *the amount of work is measured by the product of the force by the distance through which the body moves along the line of action of the force.* Thus when in digging a ditch a ton of earth is thrown to an average height of 6 ft., the work done is $2000 \times 6 = 12,000$ foot-pounds.

If the motion of the body is not in line with the resultant force, then *in estimating work only that component of the motion which is in the direction of the force is to be taken into account.* For instance, in raising a barrel into a wagon *the work done is the same whether the barrel is lifted directly from the ground or rolled up an inclined plane.* For the weight of a body is a force that acts vertically downward, consequently *in estimating work done against weight, only the vertical distance through which the body is moved is to be considered.*

When a body yields to a force work is said to be done *by the force* or *upon the body;* but when a moving body is retarded by

some resisting force, work is then said to be done *by the body* or *against the force.*

The work done in raising a weight or compressing a spring is the same whether done in a second or in an hour. The time required to do the work determines the *rate of working*, but has nothing to do with the amount of work.

It is remarkable that although *force* and *distance* are both vector quantities, *work*, which is their product, *is not a vector quantity.* It has nothing to do with *direction*, and consequently to get the total work done upon a body by several different forces, the work of each may be reckoned separately and then the sum taken.

Motion is essential to work. A great weight may rest on a support, but no work is done in supporting the weight though a great force is exerted.

65. Rate of Working. Horse-power. A given amount of work may be done either in a short time or a long time, and in commercial operations the *rate of working*, or the work done per second or per hour, is an important consideration. Thus in case of an engine we wish to know how much work it can do in a given time, and its rate of working is known as its *power.*

Power may be measured by the number of grams weight that can be raised one centimeter per second, or by the number of pounds that can be raised one foot per second; but the unit of power introduced by James Watt and commonly used in engineering practice is the *horse-power* (written H.P.).

One horse-power = 550 foot-pounds per second, or 33,000 foot-pounds per minute.

That is, a 10 H.P. engine can raise 330 lbs. through a height of 100 ft. in one-tenth of a minute, or 3300 lbs. through a height of 10 ft. in the same time.

66. Energy. The importance of the idea of work lies in the fact that a body upon which work is done acquires thereby capacity to do an equal amount of work in returning to its original state. *The capacity to do work is called energy.* Thus work is done when a spring is bent, and the spring acquires energy which is measured by the work that it can do as it unbends. Also a 10-lb. weight raised 100 ft. above the earth has had 1000 ft.-lbs. of work expended in raising it, and it has gained the power

to do that same amount of work in returning to its original position.

The energy of a bent spring resides in the spring itself in virtue of its internal stresses; but in case of the raised weight the energy belongs not to the weight alone, but to the system of two bodies, the earth and the weight, which are separated in opposition to the stress or attraction between them.

67. Kinds of Energy. In both illustrations given above the energy depends on the relative positions of bodies or parts of bodies between which there exist stresses. There is another form of energy which depends not upon stress, but upon the motion of matter.

Suppose the raised weight is set free and allowed to fall with nothing to resist it, the force of the earth's attraction is exerted upon the mass as it falls and consequently work is done and energy expended, but in this case the work is all spent in giving velocity to the falling mass. When the weight reaches the bottom it has lost all its advantage of position, but it still has power to do work in virtue of its motion; and experiment shows that the work it can do before coming to rest is exactly equal to the work that was done upon it in giving it motion. The mass, therefore, still retains the energy that it had in the raised position, but it is now *energy of motion.*

The energy which a body or system of bodies has in virtue of position or configuration is called potential energy.

The energy which a body has in consequence of the velocity of its mass is called its kinetic energy.

68. Illustration. If a mass is hung so that it can freely swing as a pendulum, when it has been raised to the position A (Fig. 33) it has been raised through the vertical distance h from B to D, and, therefore, has more potential energy at A than at B by the work done in raising it from B to A. If allowed to fall freely it will reach the bottom, moving with sufficient velocity to carry it up to C on the same level as A. At the bottom the mass has energy of motion or kinetic energy. It has entirely lost the advantage of position which it had at A, the work done in raising it to A being now wholly transformed into energy of motion.

But as the mass rises from B toward C it loses velocity for it is doing work and using up the store of kinetic energy that it received in falling, changing it again to potential energy. The pendulum has thus a constant store of energy which changes back and forth from one form to the other,

the sum of the two being always constant, except as energy is gradually lost through friction and air resistance.

69. Work against Friction. There is one case, however, in which the work done upon a body does not *seem* to increase its energy or power to do work. When a weight is pushed from one point to another on a level table force has to be exerted to overcome friction. The weight, how-ever, remains at the same level above the earth and has no more power to do work in the new position than before it was moved. The work expended seems to be quite lost.

FIG. 33

But investigation has shown (§ 417 *et seq.*) that whenever work is done against friction *heat* is de-veloped in amount exactly propor-tional to the work done; and also that when work is obtained from a heat engine a precisely cor-responding amount of heat disappears. It is, therefore, concluded that *the work which seems to be lost in friction is not really lost or annihilated, but is transformed into heat as into another form of energy.*

When, therefore, a pendulum comes to rest in consequence of friction (at its point of support, or between it and the air through which it swings) the original energy of the pendulum is not lost but transformed into heat.

FIG. 34. Prony Brake
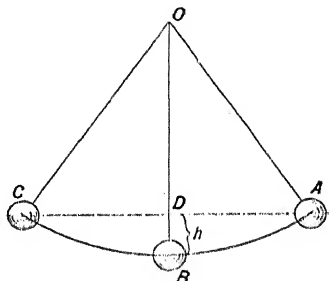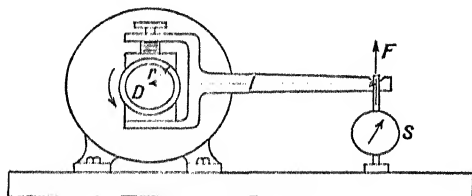
70. The Prony Brake. The prony brake is a device commonly used to measure the amount of power de-veloped by rotating machinery through work done against friction. Figure 34 shows a form of prony brake applied to a small motor. It consists of an arm one end of which is pressed against a revolving drum *D* by means of the hand screw and spring, and the other end of which is attached to a spring dynamometer *S*, by which the upward pull *F* of the arm can be measured. The revolving drum is coupled to the motor shaft by which it is driven.

The pressure on the drum is so adjusted by the hand screw [that] the drum can turn against friction, thereby causing the motor to do work [as] it revolves. A small amount of water may be put in the drum to prevent [its] becoming overheated by friction.

The work done during one revolution of the motor is evi[dently] equal to the frictional force f at the surface of the drum times the cir[cum]ference of the drum and is equal to $2\pi r f$, where r [is] the radius of the drum.

The power P, however, is the rate at which work is done (§ 65) which will be taken as the work per second in this case. Thus

$$P = 2\pi r f n \tag{1}$$

where n is the number of revolutions of the drum in one second. By the principle of the lever (§ 82), the moment of the frictional couple fr and that due at the dynamometer Fl must balance:
therefore

$$fr = Fl \text{ or } f = F\frac{l}{r}.$$

Substituting this value of f in equation (1),

$$P = 2\pi n l F.$$

If l is in feet and F in pounds,

$$P = \frac{2\pi n l F}{550} \text{ H.P.} \tag{2}$$

since 1 H.P. = 550 foot-pounds per second.

71. Forms of Energy. From the results of innumerable experiments physicists have concluded that not only is heat a form of energy, but sound, light, and all electrical and magnetic actions are manifestations of energy, and require energy to be expended in causing them, just in proportion as they are capable of doing mechanical work or developing heat.

The different manifestations of energy may be summarized as follows:

Energy of masses	Masses in motion — kinetic.	
	Elastic bodies in a state of stress.	
	Gravitation, energy of attracting masses	potential.
	Sound, both kinetic and potential.	
Energy of molecules and atoms	Heat.	
	Molecular and atomic energy.	
	Chemical action.	
Energy of ether	Electric and magnetic phenomena.	
	Light and radiation.	

When the energies involved in all these varied phenomena are studied it is found that one form of energy may be transformed into another, and that again into a third; but in every change the *amount* of the energy as measured by its ability to do work or to develop heat remains the same.

72. Conservation of Energy. The recognition of these varied forms of energy and careful measurements of the transformations from one form to another have led to the enunciation of a great principle or law known as the *Conservation of Energy*, which may be thus stated:

In any system of bodies which neither receives energy from without nor gives up any, the total amount of energy is unchanged whatever actions or changes may take place within that system, whether the energy manifests itself in mechanical forms, in sound, heat, light, electric, or magnetic effects, or in chemical action or molecular or atomic changes.

In most cases the tracing of all the changes is a difficult matter. For example, a cannon ball receives energy from the work done by the powder gases as they expand forcing the ball from the gun. As it travels it is resisted by the air, losing kinetic energy exactly equivalent to the heat energy developed by friction in the air. On striking the target, sound waves carry off a small part of the energy, there may also be a flash of light which also takes away some energy, and the rest will be found in the form of heat developed in the target and in the ball itself and also in the form of kinetic energy in the fragments which may be thrown off. The principle of the conservation of energy asserts that if we add together all the energy that is derived from the motion of the ball the sum will be exactly equal to the amount of work which was required to give it its motion.

This law is the most important and extensive generalization of the science of physics, and much of the progress of modern physics is due to its recognition. Every experiment in which the quantities of energy can be accurately determined is a test and confirmation of its truth, and no principle of physics is better established.

In consequence of this law, the determination of the energy involved in any action assumes new importance and is an essential part of the study of every physical phenomenon.

73. Availability of Energy. The presence of friction and analogous forms of resistance everywhere in nature causes a constant transformation of various forms of energy into heat, in which stage it is conducted from one body to another and gradually becomes uniformly diffused so that although the energy still exists it is no longer available for the purpose of obtaining other forms of energy that may be desired. There is thus a constant degradation of energy going on throughout the universe, more available forms being constantly frittered away into heat.

<div align="center">FRICTION</div>

74. Friction. When one body slides over another the motion is resisted by a force which is called *friction.* It is always a resistance, acting against the motion, and depends on the character of the surfaces in contact and on the force pressing them together.

It is a force of the greatest importance in daily life. If it were not for friction, nails and screws and knots would not hold, ropes could not be made, nor could we even walk across a floor. On the other hand, we would gladly be rid of friction in machines, for it is the cause of a large proportion of energy being lost in heat.

Friction appears to be due to the interlocking of minute roughnesses on the surfaces, together with the clinging together or adhesion of the points of closest contact. It is, therefore, diminished by polishing the surfaces, which diminishes the roughness and also makes the points of contact broader so that the film of air or oil is more effective in preventing adhesion.

When two surfaces have been resting in contact the friction at starting is greater than after the motion has been established. It seems probable that this may be due to the closer contact due to the film of air or oil being squeezed out by the continued pressure.

Friction also resists the rolling of one body on another, though *rolling friction* in case of two given surfaces is much less than *sliding friction.* Rolling friction when surfaces are well polished appears to be due both to cohesion and to a slight deformation both of the surface and of the roller at the point of contact, for the surface is compressed as it passes under the roller, and though

it may spring back again *it does not exert quite as much force in recovering as it opposed to the deformation*.

75. Laws of Friction. Let the block P be drawn along by the weight F, which is not sufficient to start it in motion, but will keep it moving with constant velocity when once started. The weight F is then equal to the force of friction, for it just balances it, neutralizing the resistance to the motion.

It is found in this way that *the friction between two given surfaces is proportional to the force pressing them together*. If the block P weighs 5 lbs. and if an additional weight of 5 lbs. is placed on the block, the force of friction is doubled.

It is also found that *the force of friction, within wide limits, is independent of the area of the surface of contact*. For instance, the friction

FIG. 35

of the block P is almost the same whether it slides on a narrow or a broad side, provided they are equally smooth.

The velocity with which one surface slides over the other makes little difference, the friction being appreciably the same for all moderate speeds; but the resistance to starting, or *static* friction, is greater than the friction after the motion is established.

It is evident, however, that these laws do not hold without limit. For if one surface is very small, as in case of a point resting on a plane surface, or if the pressure is so great that one body presses into the other, then one cannot move on the other without tearing or injuring the surface, and the law no longer holds.

76. Coefficient of Friction. It follows from the first law of friction that *the force of friction divided by the force pressing the surfaces together is a constant*, this constant is called *the coefficient of friction* of the surfaces concerned; *it is a fraction which when multiplied by the force pressing two surfaces together gives the force of friction to be overcome*.

Thus if the coefficient of friction in case of iron wheels on iron rails is 0.004, then, if the wheels weigh 1000 lbs., a force of 4 lbs. will be required to overcome the friction.

When an engineer wishes to know how much force will be
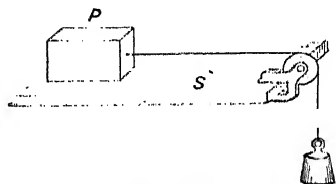
required in moving a house to cause it to slide on its ways, he has only to multiply the coefficient of friction for the soaped beams on which the house rests by the weight of the house itself.

Some Coefficients of Friction

Sliding Friction

Oak upon oak, fibers parallel { without lubricant		0 42
{ rubbed with dry soap		0 16
Oak upon oak, fibers crossed without lubricant		0 29
Iron on bronze { without lubricant		0 25
{ thoroughly lubricated, may be as small as		0 06

Rolling Friction

Cast-iron wheels on rails.	0 004

77. Limiting Angle of Repose. The angle at which a surface may be inclined before a body resting on it begins to slip down is determined by the coefficient of friction between the surfaces. Thus let a weight W rest on a surface inclined at an angle θ. The earth attracts the weight with
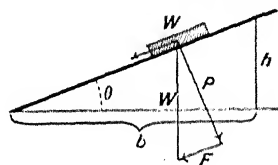
a force W which acts vertically downward. We may resolve this force into two components, one P which is perpendicular to the inclined surface and represents the pressure of the weight against the surface, and another F which is parallel to the surface and represents the force urging the weight down along the slope. The force of friction between the weight and the inclined plane

FIG. 36

is equal to the product kP, where k is the coefficient of friction. If the friction is less than F the weight will slide with increasing speed down the incline, while if it is greater than F the weight will remain at rest.

It will be noticed that P is made smaller by increasing the slope of the incline, and since k remains constant, the force of friction is less the greater the slope, and is zero when the slope is vertical.

At one particular angle θ, which may be called the limiting angle of repose, the force of friction balances the force F, and we have $kP = F$. *At that angle the weight does not start to slide of itself but if started, slides down with constant speed.* In this case $k = \dfrac{F}{P}$ and by similar triangles $\dfrac{F}{P} = \dfrac{h}{b}$, therefore $k = \dfrac{h}{b}$ or $k = tan\ \theta$.

Hence by finding the limiting angle of repose in a given case the coefficient of friction is at once determined.

78. Means of Diminishing Friction. To make friction small the surfaces should be very hard and of fine even polish. Where there is much wear it is customary to make one of the bearing surfaces of a harder material than the other. Thus the crank pins on steam engines are made of polished steel and turn in brass boxes, the friction between the brass and steel being less than it would be between two parts of steel.

Rolling friction is very much less than sliding friction, therefore, wheels are used on carriages, etc. It depends to some extent on the diameter of the wheels, being less when the diameter is greater. But even when wheels are used there is sliding friction in the hubs. The resistance to the motion of the vehicle due to this sliding friction is diminished by making the axles of small diameter, but the length of the axle in the hub of the wheel or the length of its bearing surface does not affect the frictional resistance.
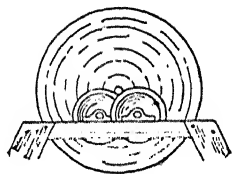
Fig. 37

To avoid the friction due to the sliding between wheel and axle, ball bearings are used; but even in these bearings there is some sliding friction where adjoining balls rub against each other.

In some cases the axle is made to rest on the rims of two smaller wheels which are called friction wheels (Fig. 37). This is a common practice in mounting grindstones.

Friction is greatly diminished by the use of *lubricants*, of which those most in use are oil, grease, soap, and black lead. The substances used as lubricants cover or wet the surfaces so that the rubbing takes place between layers of these substances instead of between the original surfaces. When the bearing surfaces are subjected to great pressure, as in heavy machinery, a thick oil is used that is not driven out by the pressure; in very light machinery, as in clocks and watches, a very thin oil is used. If oil were used on wooden bearings it would only increase the friction, for it would soak into and swell the wood; dry soap or paraffin may be used as a lubricant for wood surfaces.

MACHINES

79. Machines. *Machines are devices by which the amount or mode of application of a force is changed for the sake of gaining some practical advantage.* Simple machines, known also as the mechanical powers, are the rope and pulley, lever, wheel and axle, inclined plane, and screw. All afford interesting cases of forces in equilibrium; but they may also be discussed from the point of view of the conservation of energy, for the work done on a machine must be equal to the work done by it if there is no loss of energy in friction.

The ratio of the force exerted by a machine to the force applied is called its mechanical advantage.

80. Rope and Pulley. In all tackles where ropes are used *the tension or force is the same at every point in a continuous rope, whether it passes over pulleys or not, if there is no friction.*

FIG. 38

Let us apply this principle to a few cases. In case 1 (Fig. 38) there is only one rope and the 100-lb. weight is supported by it, therefore all parts of the rope are under a tension of 100 lbs., and that force must be exerted at *P* in whatever direction the pull may be made.
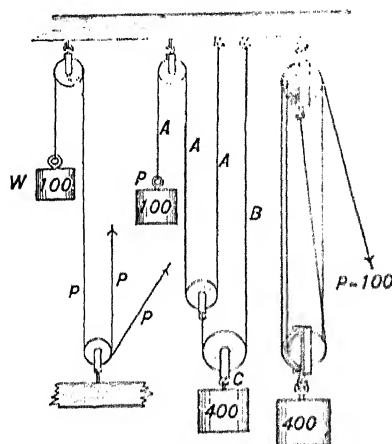
In case 2 the 100-lb. weight is supported by the rope *A*, all parts of this rope are therefore under that tension; but *B* is attached to a pulley which is drawn up by *two* parts of *A*. Since the pulley is in equilibrium it follows that the upward pull of the two parts of *A* must be equal to the downward pull of *B* together with the weight of the pulley. If we neglect the latter the tension on *B* must be 200 lbs. and similarly that on *C* must be equal to twice that on *B*. Hence, neglecting friction and the weight of the pulleys, a weight of 400 lbs. on *C* will balance a weight of 100 lbs. on *A*.

In case 3 there is one continuous rope which is fastened at the top and passes over two sheaves in each pulley, the lower pulley is therefore sustained by four parts of one rope, hence when a weight of 400 lbs. is supported by the lower pulley the tension on the rope is 100 lbs.

The mechanical advantage in the first case is 1, while in the second and third cases it is 4.

81. Principle of Work Applied. From the conservation of energy it is clear that the work done by a machine must be equal to the work done upon it, provided there is no friction and the energy stored in the machine is not changed. In illustration of this principle consider the various tackles of the preceding paragraph, and let x represent the distance that W is raised in a given case while the end of the rope at P is pulled through a distance y. Then the work done by the machine when W is raised is Wx, and the work spent in raising the weight is Py, and therefore $Wx = Py$.

In the first case $x = y$, therefore $P = W$.

In the second case $x = \frac{1}{4}y$, therefore $\frac{1}{4}W = P$.

In the third case also $x = \frac{1}{4}y$, therefore $\frac{1}{4}W = P$.

It should be noted that in the last two cases if the weights of the pulleys are taken account of we cannot say that $Wx = Py$, for some of the work done is spent in raising the movable pulleys. Thus, in case 2, if each pulley weighs w, we have

$$Py = w\frac{y}{2} + (w + W)\frac{y}{4} \text{ or } P = \frac{1}{2}w + \frac{1}{4}(w + W).$$

82. Lever. In the lever a rigid bar resting on a point of support, or *fulcrum*, is used to exert a great force near the fulcrum when a smaller force is exerted at the end of the longer arm of the lever. A crowbar as used in moving a stone, a hammer in drawing a nail, are examples of levers. Levers are sometimes divided into three classes depending on the relation between the position of the fulcrum and the points where the weight is raised and the force applied, as shown

Fig. 39 Classes of levers

in the figure, where P represents the force applied to support the weight W, and F is the fulcrum.
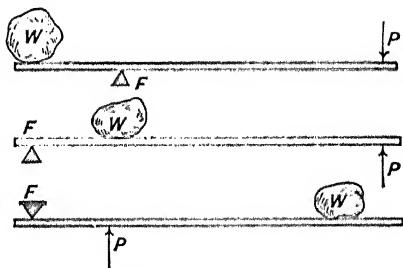
The upper lever in the figure belongs to the *first class;* the next to the *second class;* and the lowest to the *third class.*

The distance from P to the fulcrum is called the power arm and

that from W to the fulcrum is called the weight arm, and the principle of moments tells us that if these distances are measured perpendicular to the lines of action of the forces P and W, then *the product of P by the power arm is equal to the product of W by the weight arm.* In other words, the moments of the two forces about the fulcrum as axis must be equal and opposite.

The pressure F against the fulcrum, since the three forces P, W, and F must be in equilibrium, is represented by the vector necessary to form a triangle with P and W. Of course if P and W are parallel, F must be either their sum or difference, depending on circumstances.

83. Crank and Axle. In case of the crank and axle, shown in figure 40, the relation between the weight and the force applied at the crank to support it, is at once obtained from the principle that the moments of F and W about the axis must be equal, since the only motion that the system can have is one of rotation. Hence if R is the length of the crank arm and r the radius of the axle or drum on which the rope supporting W is wound, we have $FR = Wr$ in case the force F acts at right angles to R.

Here, again, we may apply the principle of work, for in one revolution of the crank the weight W is raised a distance equal to the circumference of the drum or $2\pi r$, while the balancing force F acts through a distance $2\pi R$. We have, therefore, in case of equilibrium

$$W\, 2\pi r = F\, 2\pi R \text{ or } Wr = FR.$$

FIG. 40

84. Inclined Plane. Barrels or casks are sometimes rolled up inclined planes and thus raised where they could not be directly lifted. The advantage of the inclined plane may be understood from figure 41, where W represents a weight resting on the inclined plane having length l, height h, and base b. The attraction of the earth is a force vertically downward on W, but it may be resolved as is shown into the components N at right angles to the inclined plane and F parallel to it.
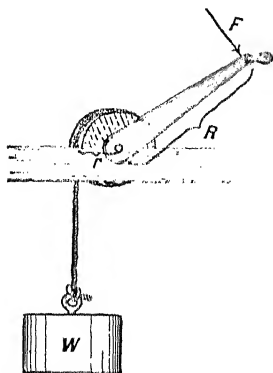
The component N is balanced by the pressure of the plane, while the component F represents the force that must be balanced by the push P necessary to support the weight on the plane. From the similarity of the two triangles it is clear that W, N, and F are proportional to l, b, and h, respectively. That is $F : W : : h : l$, or in words, *the force required to support the weight on the inclined plane is to the whole weight as the height of the plane is to its length.*

FIG. 41

The same conclusion may also be reached by the principle of work, for if the weight is pushed up the plane the supporting force P acts through the length l, while the weight W is only raised against the earth's attraction through a distance h. Hence $Pl = Wh$.

If the force P, instead of acting parallel to the *length* of the inclined plane, were parallel to its base we should resolve the weight W into components N and F as in figure 42, where F is parallel to the base. Then

$$F = P = W \frac{h}{b}.$$

85. Screw. The screw as used in the ordinary letter press may cause enormous pressures by the application of a very moderate force to the lever arm. In one complete revolution of the screw it advances the distance between consecutive threads measured parallel to the axis. This distance is called the *pitch* of the screw.

The mechanical advantage of the screw may be determined by con-
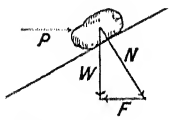
FIG. 42
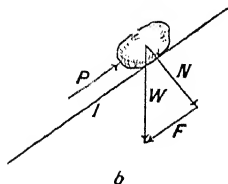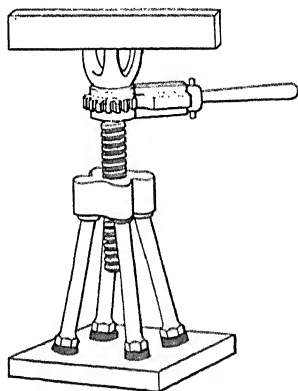
FIG. 43

sidering the thread as a sort of inclined plane wrapped around the axis, but we may deduce it more conveniently from the prin-

ciple of work; for if the force P operating the screw acts at right angles to the end of a lever arm of length R, in one revolution of the screw the force P acts through a distance $2\pi R$, while the screw advances through a distance h equal to the pitch of the screw. Hence if W is the force exerted by the screw we have by the principle of work

$$2\pi R P = W h$$

or

$$\frac{W}{P} = \frac{2\pi R}{h}.$$

86. Chinese Capstan and Differential Pulley. In the Chinese capstan a drum or axle having two parts of somewhat different diameters is operated by lever arms or capstan bars, so that one end of a rope is wound up on
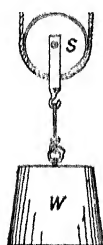
FIG. 44   FIG. 45. Differential Pulley

the drum of larger diameter while the other end unwinds from the smaller drum. The rope passes around a pulley S which is attached to the anchor or other weight to be raised. The force W is divided between the two parts of the rope pulling on S, so that the rope is under a tension $\frac{W}{2}$. If r and R are the radii of the small and large drums, respectively, the moments of

the forces exerted by the rope on the drum are $\dfrac{W}{2}r$ and $\dfrac{W}{2}R$ and the difference between these two moments must be balanced by the moment of the force P acting on the end of the capstan bar of length l. Hence we have in case of equilibrium

$$Pl = \frac{W}{2}(R - r).$$

The advantage of such an arrangement is evidently the same as if one end of the rope were fixed and the other, after passing around S, were wound up on an axle whose radius was $R - r$. But such an axle being of small diameter would not have the strength of the larger axle with two drums.

The differential pulley is a similar device used for raising heavy weights. There is an upper pulley having a single sheave with two grooves of different diameters like the two drums of the Chinese capstan. An endless chain passes over one groove in the upper pulley then around a pulley attached to the weight to be raised, and then around the second groove of the upper or fixed pulley. The grooves of the upper pulley have notches to receive the chain so that it cannot slip, and the chain is passed over it in such a way that it is wound up on one groove at the same time that it unwinds from the other. If the difference in diameters of the two grooves in the upper sheave is small, a small pull on the chain may suffice to support a large weight.

PROBLEMS

1. A 180-lb. barrel is rolled up an inclined plane 12 ft. long to a platform 4 ft. above the ground. How much force must be exerted along the plane and how much work is done? Find also the force and work when the plane is 20 ft. long, the height being the same.

2. Find the force which the barrel exerts against the plane in both the cases specified in the first problem.

3. How much force parallel to the plane is required to support a weight of 39 kgms. on a frictionless inclined plane 13 meters long and 5 meters high? Also find the force with which the weight presses against the plane.

4. If the coefficient of friction between weight and plane in the last question is 0.20, find the force of friction and how much force must be exerted parallel to the plane in drawing the weight up, also in lowering it.

5. When the coefficient of friction between a weight and the inclined plane on which it rests is 0.30, find the ratio of its height to length when the plane is so steep that when the weight is started it slides down without acceleration.

6. A certain jack-screw has a screw 2 in. in diameter with three threads to the inch, and is operated by a lever arm 2 ft. long. What weight can be raised by a force of 48 lbs. applied at right angles to the end of the lever arm, neglecting friction?

7. When the coefficient of friction of the oiled surfaces of the jack screw described in problem 6 is 0.06 and when a weight of 5 tons is raised, find the force required at the end of the lever arm to overcome friction, and the additional force required to raise the weight.

8. In problem 7, find the ratio of the work required to raise the weight 1 ft. without friction, to the actual work with friction, and thus determine the efficiency of the screw. Would the efficiency be the same if one half as large a weight were being raised?

9. Find the tension on a bicycle chain when the pedal is pressed down with a force of 120 lbs.; the crank arm being 6 in. long and the sprocket wheel 8 in. in diameter.

10. If a force of 40 lbs. must be exerted on the arm of a windlass in raising a weight of 120 lbs. while a force of only 20 lbs. is required in lowering the same, find the force expended in overcoming friction, and the efficiency of the windlass, and what per cent of the work done is lost in friction.

11. How much force must be exerted on the crank of a windlass to raise a weight of 180 lbs., if the crank arm is 20 in. long and the drum on which the rope is wound is 8 in. in diameter.

12. Find the direction and amount of the force on the bearings of the windlass in the previous question, first, when the crank is in a horizontal position and being pressed down; second, when the crank arm is vertical.

13. A man weighing 150 lbs. raises himself in a sling by means of a rope passing over a movable pulley attached to the sling and a fixed pulley overhead. With how much force must he pull? Show also how to obtain your result by the principle of work.

14. A man weighing 180 lbs. runs up 24 steps, each 7 in. high, in 8 seconds. How much work does he do and what horse power does he expend?

15. A donkey-engine is required to raise by means of a tackle a 2 ton weight to a height of 100 ft. in ½ minute. What horse power is required if the efficiency of the tackle is 70 per cent?

16. When 1 H.P. is expended by a horse in pulling a load at the rate of 6 miles per hour, find the force with which the horse pulls the load.

17. What load can two horses draw along a level road at the rate of 3 miles an hour if they spend 2 H.P. in pulling the load, when the coefficient of friction of wagon on road is $\frac{1}{10}$. Ans. 2500 lbs.

18. A locomotive drawing a train along a level track at 30 miles per hour expends 75 H.P., find the total air and frictional resistance overcome.

Ans. 937.5 lbs.

19. A locomotive draws a 300-ton train along a level track at the rate of 20 miles per hour; while working at the same rate it draws it up a ¼ per cent grade at 15 miles per hour; what horse power is expended, supposing the frictional and air resistances the same in both cases, and what is the resistance in pounds. **Ans. Resistance = 4500 lbs.; H.P. = 240.**

III. KINETICS OF A PARTICLE

Rectilinear Motion of a Mass

87. Introductory. Up to this point we have studied especially cases of *equilibrium*, where the forces acting are so balanced that there is no acceleration. We must now examine in some detail the various forms of motion where forces are involved in such a way as to cause acceleration.

This part of mechanics, as Mach says, "is a wholly modern science. All that the Greeks achieved in mechanics belongs to the realm of *statics*. *Dynamics* was first founded by Galileo."

Before 1638, when Galileo first published the results of his experiments, so little progress had been made in this direction that it was currently held that heavy bodies fell faster than light ones.

In studying the effect of force in giving motion to matter, the simplest case to examine is where a definite portion of matter is acted on by a constant force. This is the case with falling bodies; for while a body is falling freely it is being urged downward by a constant force which we call its weight. Therefore, Galileo carefully studied the motion of falling bodies, and of bodies rolling down inclined planes, and showed that in each of these cases the motion was with *constant acceleration*. As pendulum clocks had not been invented at that time, he made use of a simple water clock to measure short intervals of time in his experiments. This consisted of a large vessel of water having a jet closed by the finger, from which water was allowed to escape during the time interval to be measured. Thus the weight of water escaping while a body rolled down an inclined plane served to measure the time of descent.

These experiments also showed that when a plane was inclined at such an angle that the force parallel to the plane required to keep a body from sliding down was one-half the weight of the body, then its acceleration in sliding down was one-half its acceleration when falling vertically. That is, *the acceleration was proportional to the force causing the motion.*

88. Atwood's Machine. A convenient device for studying the effect of forces in giving motion to masses is the apparatus known as Atwood's

machine (Fig. 46). Two equal weights *A* and *B* are hung over a very light, carefully balanced wheel mounted so that it will run with as little friction as possible. An additional weight or rider *w*, having two projecting arms, is laid on top of the weight *A*, which is supported so that it can be liberated at any instant. When the weight *A* is freed it moves down, accelerated by the rider *w*, until it reaches the ring *C* which picks off the rider *w* and allows *A* to pass freely through. After passing the ring *C* there is no longer any accelerating force, since the rider is removed, and the weight *A* continues to move with the velocity which the rider had given to it.

Fig. 46　Atwood's machine

Thus if the ring *C* is so adjusted that *A* passes through it exactly 2 seconds after being liberated, and if *D* is so placed that *A* moves from *C* to *D* in the next second, then if *C* and *D* are found to be 30 cm. apart, we conclude that *A* acquired a velocity of 30 cm. per second by a force which acted steadily for 2 seconds. If the same force is now allowed to act for 1 second, a velocity of only 15 cm. per sec. will be acquired. By varying the weight of the rider or using instead of *A* and *B* a pair of weights, having double the mass, the following conclusions may be established:

(*a*) The motion is with constant acceleration.

(*b*) The acceleration is proportional to the weight of the rider so long as the total mass *A* + *B* + *w* is constant.

(*c*) If the mass of the moving system is doubled, a given rider will cause only half as great acceleration as before.

89. General Principle. The effect of a force in giving motion to a body, as brought out in the experiments just described, may be thought of as du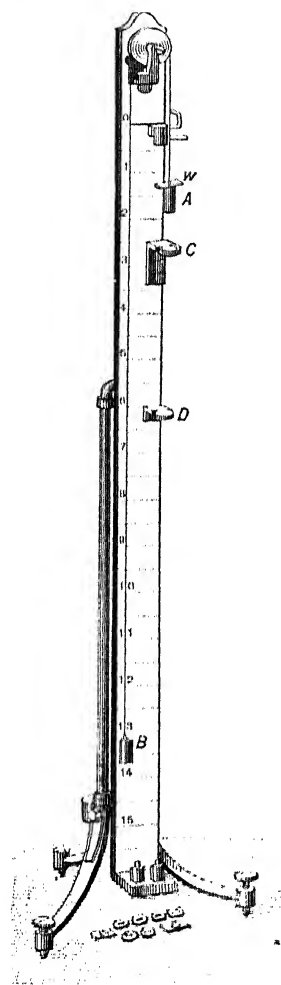e to a general principle which may be thus stated: *the effect of a force in changing the motion of a mass is not in any way affected by the state of rest or motion of the mass which is acted upon.*

For instance, while a force is acting on a mass and increasing its velocity, suppose a second and equal force to act in the same direction upon the same mass. The second force being equal to the first will produce just as great an increase in velocity per second as is being produced by the first; and since both effects take place simultaneously and without interference, the total change in velocity will be twice that which would have been produced by the original force. It follows that *the change in velocity per second when a force acts on a body is proportional to the amount of the force.*

And the change in velocity of a body when acted on by a force is also *proportional to the length of time during which the force acts*, for suppose a mass has acquired velocity by a force acting upon it for one second, if the force now acts for another second it will increase the velocity of the mass as much more in the same direction, since the effect of a force is in no way conditioned by the state of rest or motion of the body upon which it acts.

90. Impulse. The change in velocity which a given mass experiences is proportional therefore both to the amount of the force and to the time during which it acts. A large force acting for a short time may produce the same change in the velocity of a mass as a small force acting for a longer time.

A billiard ball may be made to roll as fast by pushing it as by striking it with the cue; the force in the second case is very much greater than in the first, but is exerted during an exceedingly short time; the impulse in both cases must be the same.

The product of the amount of a force by the time during which it acts is called the *impulse*.

91. Force and Motion. Again, suppose two equal masses moving side by side are acted on by equal forces in the same direction, they will both gain in velocity equally and will accordingly continue to move side by side, and their motion will evidently not be affected in any way if the two masses are connected forming a single large mass.*

From this consideration we see that if a force gives a certain acceleration to a given mass then twice the force will be required

* This cannot be regarded as known *a priori* for it results from the experimental fact that the inertia of one body is not affected by its proximity to another.

to give the same acceleration to a mass twice as great, etc. Or, in order that different masses may all have the same change in velocity per second, the forces acting on them must be proportional to the masses.

But if the mass is doubled without any corresponding change in the force which acts upon it, the gain in velocity will be only half as great as before, for the motion in that case will be the same as if the original mass were acted on by half the original force.

92. Momentum. A given impulse may produce a great change in the velocity of a small mass, or a proportionally small change in the velocity of a greater mass; therefore, to measure the effect of an impulse, a quantity is employed which is proportional both to the mass and velocity of the moving body; this is called its *momentum.*

The momentum of a body is the product of the amount of its mass by the amount of its velocity, and is a directed or vector quantity.

93. Three Laws of Motion. The relations between forces, masses and motion, were first clearly enunciated in the form of three laws of motion by Sir Isaac Newton in his celebrated *Principia,* published in 1686. Two of these laws have been already discussed (§§ 31, 38), but are here repeated in order that all three may be presented together.

First Law: Every body continues in its state of rest or of moving with constant velocity in a straight line, unless acted upon by some external force.

Second Law: Change of momentum is proportional to the force and to the time during which it acts, and is in the same direction as the force.

Third Law: To every action there is an equal and opposite reaction.

94. Discussion of Second Law. This law may be also expressed in the formula

$$mv - mu \propto Ft$$

where F is a force acting on a mass m for a time t, and u is the velocity at the beginning of the time interval t, while v is its velocity at the end of that time. Thus mv is the momentum after the force has acted, while mu is the original momentum of the mass. The gain in momentum is, therefore, $mv - mu$, and ac-

cording to the law this is proportional to the force F and to the time t jointly, or to their product Ft.

The above formula may be written:

$$Ft = k(mv - mu) \text{ or } F = km\left(\frac{v - u}{t}\right)$$

where k is constant, the value of which depends on the particular units which are employed in measuring the various quantities concerned.

In the above equations F represents the *average* value of the force during the time t in which the velocity of the mass has changed from u to v; but when t is exceedingly short, $\dfrac{v - u}{t}$ approaches as its limit the actual rate of acceleration at the given instant, while F is the corresponding force at that same instant, and we may write,

$$F = kma \tag{1}$$

that is, *the acceleration of a body is proportional to the force acting upon it and inversely proportional to its mass.*

This may be called the fundamental formula of dynamics as it is a direct expression of the second law of motion, is absolutely general, and enables us to determine the forces acting in any case where the mass and motion of a body are known, since the acceleration is determined from the motion.

Thus it follows that if the force acting on a mass is constant the mass moves with constant acceleration, while if the force varies the acceleration varies in the same proportion.

95. Dyne and Poundal. In dealing with cases of equilibrium we have used the ordinary gravitation measures of force, the weight of a pound or gram, but in studying the accelerating effect of forces it will be found more convenient to use as the unit a force which will make the constant k equal to unity in the above expression, so that we may write simply

$$F = ma$$

Defined in this way, *unit force is one which will give unit acceleration to unit mass,* or *unit force acting for unit time on unit mass will change its velocity by unity.*

When the centimeter gram and second are the fundamental units as in the C. G. S. system, the unit force is called the *dyne*, from the Greek word for force. It is *a force which, acting on a mass of one gram for one second, will change its velocity by one centimeter per second*.

Hence to find the force in dynes in a given case of motion it is only necessary to multiply the mass in grams by its rate of acceleration measured in centimeters per second per second.

Thus a force of 100 dynes acting on a mass of 10 grams will give it an acceleration 10, or in one second will give it an increase in velocity of 10 cms, per second.

A unit of force similarly based on the foot, pound, and second as units of length, mass, and time, respectively, is sometimes used and is called the *poundal*, it is *the force which acting on a mass of one pound will increase its velocity one foot per second for every second that it acts*.

The dyne and poundal have the advantage of being absolutely definite units of force, and *do not vary from point to point on the earth as the weight of a gram or pound varies*.

96. Unit of Work or Energy. The unit of work on the C. G. S. system of units where the force is measured in *dynes* and the distance in centimeters is known as the *erg* from the Greek word for work. *It is the work done when a body moves one centimeter in the direction in which it is urged by a force of one dyne.*

The corresponding unit of work or energy on the foot pound-second system is the *foot poundal*, and *is the work done when a body moves one foot in the direction in which it is urged by a force of one poundal*.

97. Motion in a Straight Line with Constant Velocity. When a body moves in a straight line with constant velocity the acceleration is zero and therefore the force must be zero according to the formula $F = ma$.

The moving mass is, therefore, in equilibrium. This is the case considered in Newton's first law of motion.

A railway train while running at constant speed is in a state of equilibrium. The force of the locomotive urging it on is exactly balanced by the resistance of the air and friction of the wheels. So when a bucket is drawn up out of a well with constant speed it

is in equilibrium and the upward pull on the rope is exactly equal to the weight of the bucket of water.

98. Motion in a Straight Line with Constant Acceleration. When a body moves in a straight line with a velocity which is increasing or diminishing at a constant rate, it has a constant acceleration in the direction of the motion in one case and opposite to it in the other.

When the acceleration a is constant, the change in velocity of the moving body in t seconds is at. And if the velocity at the beginning of the time t is u, and that at the end of the time is v, then

$$v = u + at \text{ when the speed is increasing;}$$
$$v = u - at \text{ when the speed is decreasing.} \qquad (1)$$

The space passed over in t seconds will be found by multiplying the *average velocity* during the interval by the time t. But since the acceleration is constant the velocity increases uniformly with the time, and therefore the average velocity is the arithmetical mean of the initial and final velocities, or $\dfrac{v + u}{2}$. The space traversed in time t, therefore, may be expressed by the formula

$$s = \frac{v + u}{2}t. \qquad (2)$$

Substituting for v its value

we find

$$s = ut \pm \tfrac{1}{2}at^2. \qquad (3)$$

But equation (1) may be put in the form

$$a = \frac{v - u}{t},$$

and this multiplied by (2) gives

$$2as = v^2 - u^2. \qquad (4)$$

By the use of these formulas (1 to 4) any two of the quantities u, v, a, t, s may be determined when the other three are given.

The student should thoroughly memorize these formulas and exercise himself in applying them to simple problems, such as those on page 72.

99. Force Causing Rectilinear Motion with Constant Acceleration. The kind of motion just discussed is produced whenever a mass is acted on by a constant force in one direction; for in such a case the acceleration is constant and given by the relation

$$a = \frac{F}{m}.$$

Thus when a car is drawn along a track by a stretched spring which is kept constantly at the same tension, the motion is with constant acceleration. So also a falling body has this kind of motion, for it is constantly urged downward by its own weight, which is a nearly constant force. When a body slides down an inclined plane, the force urging it down along the plane is the same at one point as at another, and, therefore, in this case also the acceleration is constant.

100. Falling Bodies. Freely falling bodies are the most familiar examples of bodies moving with constant acceleration. For a body near the surface of the earth is attracted or urged downward with a certain constant force which we call its weight, and when it is set free so that its weight is the only force acting, it falls with constantly accelerated motion. In ordinary experience, however, where bodies fall through air, the resistance of the air is another force which modifies the motion. If the resistance in a given case were constant, the body would still fall with constant acceleration, but the air resistance increases greatly with the velocity of the falling body, so that in case of a light body, as the speed increases the air resistance may become equal and opposite to its weight, and when that is the case it falls without acceleration. This is the case with scraps of paper and rain drops.

Strictly speaking, even the weight of a body is not constant as it falls, but increases as it approaches the surface of the earth. The weight of a kilogram one mile above the earth's surface is less by $\frac{1}{2}$ a gram than at sea level, and at the ceiling of a room 3 meters high a kilogram weighs about one milligram less than at the floor. This variation of force with height causes a corresponding increase in the acceleration of a falling body as it approaches the earth's surface; but this is so small, however, that except in case of great heights it may be neglected.

101. Acceleration of Gravity. The early philosophers speculated as to *why* bodies fell; Galileo was the first to carefully determine *how* bodies fell. He also showed, contrary to the universally accepted opinion of his day, that except for air resistance all falling bodies are equally accelerated. A large stone or a small one, an iron cannon ball, a lump of lead, or block of wood when dropped from the top of a tower reach the ground in the same time. If a feather, scraps of paper, and some bits of metal or lead shot are placed in a long tube (Fig. 47) from which the air is exhausted, on quickly inverting the tube all reach the bottom at the same instant. Hence the rate of increase in velocity, or *acceleration, is constant* at any given place on the earth *for all kinds and sizes of bodies.*

This constant acceleration is called the *acceleration of gravity* at the given place, it is usually represented by the symbol g and is measured most accurately by pendulum experiments.

The value of g at the sea level for the latitude of New York is 980.2 cm./sec.², or 32.16 ft./sec.² The table on page 117 gives the values at some other places.

The formulas for falling bodies are, therefore, obtained from those of § 98 by making the acceleration equal to g. Thus

$$v = u + gt$$
$$s = ut + \tfrac{1}{2} gt^2$$
$$2gs = v^2 - u^2.$$
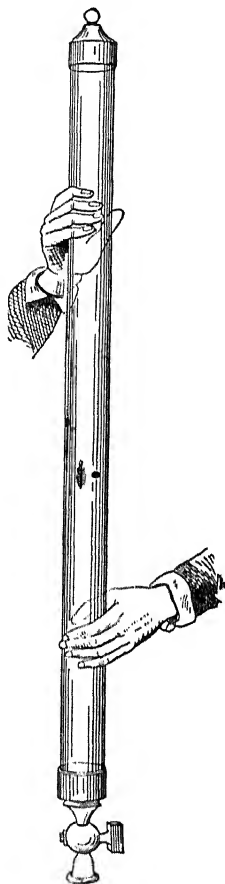

FIG. 47   Fall in vacuo

When a body is simply dropped, with no initial velocity, u is zero, and we have

$$v = gt$$
$$s = \tfrac{1}{2} gt^2$$
$$2gs = v^2.$$

In approximate calculations and in working problems for practice, g may be taken as 980 cm. sec.² or 32. ft. sec.²

102. Mass Proportional to Weight. Galileo's discovery that all kinds and sizes of bodies when dropped to the earth at the same place are accelerated at the same rate except for air resistance, leads to an important conclusion. For when two bodies are equally accelerated their masses must be proportional to the accelerating forces (§ 91), which forces, in the case under consideration, are the weights of the bodies; therefore *the masses of bodies are proportional to their weights, if weighed at the same place.*

103. Relation between Dyne and Gram. The force urging downward a freely falling mass m is expressed by the formula

$$F = mg,$$

the force being in *dynes* if C. G. S. units are used. Suppose $m = 1$ gram and $g = 980$, then $F = 980$ dynes; but the force with which a mass of one gram is attracted toward the earth is called the weight of one gram, therefore the weight of one gram $= 980$ dynes, or the force which we have called a dyne is slightly more than the weight of a milligram at the earth's surface.

The student may show similarly that one poundal is about equal to the weight of a half-ounce; that is, *one pound weight at New York* = 32.16 *poundals.*

104. Gravitation Units of Force. The weight of a gram or pound is often a convenient unit of force; indeed, engineers in English speaking countries almost always measure forces in pounds; for though the weight of a pound varies from place to place on the earth, its weight at some selected spot may be taken as standard.

For example *the standard force of a pound* may be defined as the weight of a pound mass at New York where the acceleration of gravity is 32.16 *ft./sec.²*, and in that case it will be equal to 32.16 poundals. So also *the standard force of a gram* might be defined as the weight of a gram at a point where $g = 980$ *cm./sec.²*, in which case it is equal to exactly 980 dynes.

If these *gravitation units of force* are used the constant k in formula (1) § 94 is no longer unity, but we have

$$(F, \text{ in pounds}) = \frac{1}{32.16} (m, \text{ in pounds}) \times (a, \text{ in ft./sec.}^2)$$

or

$$(F, \text{ in grams}) = \frac{1}{980}(m, \text{ in grams}) \times (a, \text{ in cm./sec.}^2).$$

But most of the formulas in this book are based on the relation $F = ma$; it will therefore be best for the student in working problems to use consistently either the centimeter-gram-second system with the force in *dynes*, or the foot-pound-second system with the force in *poundals*, changing, when required, dynes or poundals into grams or pounds weight by dividing by 980 or 32 as the case may be; 32 being used instead of 32.16, as the value of g in ft./sec.², for convenience in numerical work.

105. Atwood's Machine Problem. Suppose two masses, one 40 and the other 50 grams, are connected by a cord running over a light frictionless pulley as in Atwood's machine, and suppose for simplicity that the mass of the cord and of the pulley may be neglected. It is required to find the acceleration and the tension on the cord.

FIG. 48

In this case the whole mass $40 + 50$ moves together and the resultant force which gives it motion is the weight of $50 - 40 = 10$ grams, or $10g$ dynes.

Since   force = mass $\times$ acceleration
we have $10g = 90 \times a$, therefore $a = \frac{1}{9}g$,
hence the acceleration is one-ninth that of a freely falling body.

This result may also be reached by considering that a force of 10 grams acting on a mass of 10 grams gives it an acceleration g, and therefore if that same force act on a mass 9 times as great it will give it an acceleration $\frac{1}{9}g$.

To find the tension on the cord consider the forces acting on the mass 40. It is urged downward by its own weight, 40 grams, and upward by the tension of the cord, which we may call T grams. It moves upward with an acceleration $\frac{1}{9}g$, as has been shown, hence the resultant force must be upward and equal to $(T - 40)$ grams or $(T - 40)g$ dynes, and we have, since $F = ma$,

$$(T - 40)g = 40 \times \tfrac{1}{9}g$$

whence

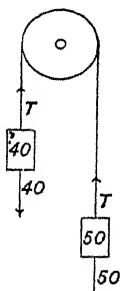$$T = 44\tfrac{4}{9} \text{ grams' weight.}$$

106. Motion on an Inclined Plane. When a mass M rests on an inclined plane, the force due to gravity, or its weight, may be resolved into two components, as shown in figure 49, one N perpendicular to the plane and the other F parallel to it. If M
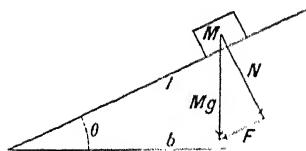
FIG. 49

is the mass in grams, its weight in dynes is Mg. And from the similarity of the two triangles, we have Mg, N, and F respectively proportional to the sides of the large triangle formed by l, b, and h. That is, $F : Mg : : h : l$ or F

$Mg\dfrac{h}{l}$ dynes. *Thus the force F causing the motion is constant, and is the same fractional part of the whole weight of the body as the height of the inclined plane is of its length. The acceleration is therefore constant.*

Since $F = Ma$, we have $a = g\dfrac{h}{l}$ or $a = g \sin \theta$.

To find the velocity which the body acquires in sliding the length of the plane l, we have only to use the formula (4) of § 98,

$$2as = v^2 - u^2.$$

The body starts from rest, hence $u = o$ and $s = l$ in this case, therefore

$$2g\frac{h}{l}l = v^2 \quad \text{or} \quad v^2 = 2gh;$$

but this is precisely the velocity which a freely falling body will gain in falling through a vertical distance h, and there is nothing in the result which depends on the slope of the plane, therefore *the velocity gained by a body in sliding down a frictionless inclined plane of any slope whatever is the same as that gained by a body in falling freely the same vertical distance.*

FIG. 50

Since the *velocity* does not depend on the slope of the plane, it will be the same at B (Fig. 50) for any smooth, frictionless

curve down which it may slide from A, and it will be the same at B as at C or D.

The *time of descent*, however, from A to B depends on the curve and may be proved to be a minimum when A and B are joined by the arc of a cycloid.

NOTE: In this section and in some of those that follow the term *velocity* should be replaced by the term *speed* according to strict usage, where evidently reference is to magnitude of the velocity without regard to its direction. This rather loose use of the term velocity, however, is common in physics. See § 23.

107. Kinetic Energy. We will now calculate the effect of a certain amount of work in giving motion to a mass m. Suppose a force of F dynes acts on m in the direction of its motion while it is moving through a space of s centimeters; the work done is by definition, Fs dyne-centimeters or *ergs*. But while the constant force F acts there is a constant acceleration a and the equations of § 98 therefore apply to the motion, and we have

$$2as = v^2 - u^2,$$

also

$$F = ma.$$

Multiplying these equations together we obtain

$$Fs = \tfrac{1}{2}mv^2 - \tfrac{1}{2}mu^2. \tag{1}$$

The change from $\tfrac{1}{2}mu^2$ to $\tfrac{1}{2}mv^2$ therefore expresses the amount of work required to change the velocity of the mass m from u to v. Starting from *rest*, the energy required to give it velocity v is $\tfrac{1}{2}mv^2$; this is also the measure of the work that the body can do before coming to rest again, therefore *the quantity $\tfrac{1}{2}mv^2$ is the measure of the kinetic energy or energy of motion possessed by a mass m moving with velocity v.*

$\tfrac{1}{2}mv^2$ $= $ kinetic energy in *ergs* when $\begin{cases} m \text{ is in grams,} \\ v \text{ is in cms. per sec.} \end{cases}$

$= $ kinetic energy in *foot-poundals* when $\begin{cases} m \text{ is in pounds} \\ v \text{ is in ft. per sec.} \end{cases}$

108. Velocity at Foot of Inclined Plane. The principle of the conservation of energy may be applied to motion on an inclined plane and leads at once to the conclusion previously stated,

§ 106, that the velocity of a body at the foot of an inclined plane depends only on its height and is independent of the slope.

For the work done in lifting the body from the bottom of the plane to the top depends only on the height of the plane, since the work is done only against gravity and serves to increase the potential energy of the body. In sliding down the plane, if no work is done against friction, all the potential energy gained will be transformed into kinetic energy, so that when it reaches the bottom its kinetic energy will be equal to the work that was done in lifting it. The kinetic energy of the body and consequently its velocity therefore will be independent of the slope of the plane.

The work done in lifting the mass m the height of the plane h is mgh, for mg is the weight of the mass expressed in dynes. The kinetic energy of the mass at the bottom is $\frac{1}{2}mv^2$, hence

$$mgh = \frac{1}{2}mv^2 \text{ and } v^2 = 2gh.$$

109. Kinetic Energy and Momentum Compared. Kinetic energy and momentum are both quantities that depend on the mass and velocity of the moving body, but while kinetic energy is expressed by $\frac{1}{2}mv^2$ and measures the *work* done on the body in giving it motion, momentum, expressed by mv, measures the *impulse* given to it, or the product of the force by the *time* during which it was acting on the body, for the second law of motion (§ 94) gives the relation

$$Ft = mv - mu,$$

or change in momentum is equal to the impulse when the force is measured in the appropriate unit.

Hence if a force acts upon a body through a certain *distance s*, increasing its velocity from u to v, and it is required to find this change in velocity, the formula for kinetic energy must be used. The kinetic energy is increased from $\frac{1}{2}mu^2$ to $\frac{1}{2}mv^2$. This change in kinetic energy is equal to the work Fs done upon the body, or

$$Fs = \frac{1}{2}mv^2 - \frac{1}{2}mu^2.$$

If, however, the *time* of the force is given, the change in velocity is found from the equation of momentum,

$$Ft = mv - mu.$$

110. Impact. When one freely moving body strikes against another there is said to be *impact.*

When the ball B is at rest and A is allowed to swing against it, if the bodies are inelastic like two balls of lead or putty, they will keep together after impact, the forward momentum of the combined mass being equal to the momentum of A before impact. If the two balls are perfectly elastic or resilient and of equal masses, like two ivory billiard balls, A will come to rest giving up its whole momentum to B, which will, therefore, swing out just as far as A has fallen. If the masses are elastic but not equal, then A may continue forward or have its motion reversed at the instant of impact depending on whether B is the less or greater mass.
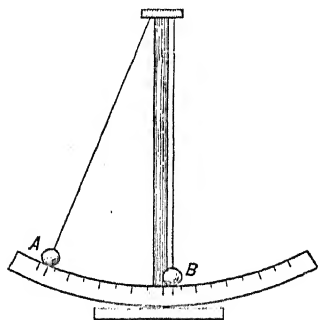
Fig. 51

In all cases of impact, whether the masses are elastic or inelastic, the total momentum of the two bodies is not changed by the impact. That is, if one body *loses* forward momentum the other *gains* an exactly equal forward momentum.

Stated algebraically,

$$Av + Bu = AV + BU$$

where A and B are the two masses, respectively, while v and u are their velocities before impact, and V and U are their velocities after impact.

This law is easily seen to be a direct consequence of the laws of motion. For at each instant during impact the forward pressure of A upon B is equal to the backward pressure of B against A, as expressed in the statement that action and reaction are equal and opposite. Hence the total forward *impulse* given to B is equal to the backward impulse sustained by A, and by Newton's second law the change in the momentum of A must be equal and opposite to the change in the momentum of B, consequently the sum of the momenta of the two is not changed.

If the two are inelastic they move together after impact with a common velocity x, whence

$$Av + Bu = (A + B)x.$$

In case of elastic bodies there is a certain instant during the impact when the compression is a maximum and the two bodies are neither approaching nor receding from each other. At that instant they are moving with the same velocity x which they would have acquired if quite inelastic. But suppose they are perfectly resilient and the pressure between them at any instant as they spring apart is exactly
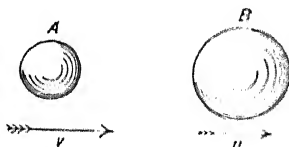
FIG. 52

equal to what it was during the corresponding instant of compression. The total backward impulse given to A will then be twice what it would have been if the bodies had been inelastic, hence the total change in velocity of A will be twice as great as $v - x$, or $2(v - x)$, and its final velocity V will be

$$V = v - 2(v - x) = v + 2x - 2v,$$

so also

$$U = u + 2x - u.$$

If the above expressions are written

$$V = v + \mu(x - v)$$
$$U = u + \mu(x - u)$$

the coefficient μ will serve to indicate the degree of resiliency. If $\mu = 1$ the bodies are quite inelastic for $V = x$ and $U = x$, but if $\mu = 2$, the resiliency is perfect.

PROBLEMS

1. A ball is thrown vertically upward with a velocity of 64 ft. per sec.; how soon will it reach the ground again and how high will it rise, and what will be its velocity when halfway up?

2. A falling body has a velocity 200 cm./sec.; how far will it drop before its velocity becomes 10,000 cm./sec.? Take $g = 980$.

3. A weight thrown forward on ice with velocity 60 ft. per sec. is resisted by a constant force, and after 5 seconds has half its original velocity; how far has it gone in that time?

4. Find the acceleration in the previous problem, also how far the weight will go before coming to rest.

5. A mass of 10 gms. is acted on by a constant force which changes its velocity from 100 to 500 cms. per sec. in 5 seconds. Find the acceleration and amount of the force.

6. What steady forward pull must be exerted by a locomotive in starting a 200-ton train to give it a velocity of 20 miles per hour in 5 minutes, neglecting friction. Find force in poundals and then in pounds.

7. A weight of 10 lbs. is thrown forward on ice with a velocity 50 ft. per sec.; if the coefficient of friction between it and the ice is 0.10, how far will it go and in how many seconds will it stop?

8. A 300-lb. mass is lowered by a rope with uniform velocity. What is the tension on the rope? If it is lowered with a constant acceleration of 10 ft. per sec. per sec. what is the tension? What if it is lowered with acceleration g?

9. An elevator weighing 2000 lbs. is pulled upward with a force of 3000 lbs. What is its acceleration, and how long will it take to gain an upward velocity of 2 ft. per sec.?

10. A mine bucket weighing 2000 lbs. and being lowered with a velocity of 3 ft. per sec. is stopped in a distance of 1 ft. What is the average force on the supporting cable while stopping?

11. If a man weighing 75 kgms. is in an elevator which is going up with constant velocity, how much force does he exert on its floor? What if the elevator has an upward acceleration of 3 meter/sec.^{2}?

12. What is the least acceleration with which a man weighing 150 lbs. can slide down a fire escape rope which can only sustain a weight of 100 lbs.? And what velocity will he have after sliding 50 ft.?

13. A 30-gm. weight is drawn up by a 70-gm. weight by means of a cord over a frictionless pulley. Find the acceleration (taking $g = 980$) and also the tension on the cord. How far will the weights move in 3 seconds from the start?

14. A 38-lb. weight resting on a level, frictionless table is drawn along by a 4-lb. weight by means of a cord over a frictionless pulley. Find the acceleration and also the tension on the cord.

15. If in the previous problem the friction between the weight and table is a force of 2 lbs. find acceleration and tension as before.

16. How many foot-poundals of work are required to give a 500-lb. shell a velocity of 2000 ft. per sec.? Find the work also in foot-pounds. If this

work is done by the powder gas in a gun 25 ft. long, and the average force in pounds against the shell as it is discharged.

17. How much energy in foot pounds must be expended in giving a 300-ton train a velocity of 30 miles an hour? If the locomotive works at the rate of 100 H.P., how long will it take to bring the train up to speed?

18. A 3-kgm. hammer with a velocity of 5 meters per sec. drives a nail 4 cms. into a plank. Find the average resistance in kilograms and how much weight resting on the nail would be required to force it into the wood.

19. A bullet weighing 1 oz. and having a velocity of 1000 ft. per sec. is fired through a plank 3 in. thick which resists it with a force of 500 lbs. With what velocity will it come out, and how many such planks would it pierce?

20. A bullet weighing 1 oz. is shot into a suspended block of wood weighing 18 lbs. 11 oz. and gives it a velocity of 6 ft. per sec. What is the combined momentum of block and bullet after impact? What was the momentum of the bullet before impact? Thence find velocity of bullet before impact.

21. What was the kinetic energy of the bullet in problem 20 before impact? What is the kinetic energy of the block containing bullet after impact? How much energy in foot pounds was expended by the bullet in penetrating into the block? What proportional part of the original energy of the bullet remains as energy of motion after impact?

22. A bullet weighing 15 gms. is shot into a suspended block of wood weighing 2985 gms. and gives it a velocity of 200 cms. per sec. Find the velocity of the bullet.

23. How high above its original level will the suspended block, in the last question, swing in consequence of the velocity given to it?

24. If all the energy of a 640 lb. shell having a velocity of 2000 ft. per sec. could be spent in raising a 10,000-ton battleship, how high would it lift it?

25. A bullet weighing 10 gms. has a velocity of 600 meters per sec. and penetrates 30 cms. into a pine log. What is the force in kilograms with which the bullet is resisted, and how far would it penetrate if at half the original velocity?

26. A monkey clings to one end of a rope passing over a frictionless pulley, and is balanced by an exactly equal weight on the other end of the rope. Explain what will happen to the counterpoise if the monkey climbs 10 ft. up the rope and then suddenly stops. The mass of the rope and wheel are to be neglected.

27. A cord passes over two fixed pulleys and hangs down vertically between them supporting a movable pulley which with attached weight weighs 5 lbs. A 3-lb. weight is hung on one end of the cord and a 4 lb.

weight on the other end. Find the accelerations of all three weights and the tension on the cord.

Note: First find a simple relation between the accelerations of the three masses from the fact that the cord is inextensible.

Motion of a Particle in Curved Path

111. Motion of a Projectile. When a body near the surface of the earth is thrown in any direction, such as AB, it is subject to the steady force of the earth's attraction vertically down-ward and, therefore, it has constantly the downward acceleration of gravity g. The initial impulse, however, gave it a forward velocity V in the direction AB, in which direction it would have continued to move with constant velocity if no force had acted on it. The actual path in which it moves may then be regarded as the resultant of motion with constant velocity V in the direction AB combined with a motion downward with constant acceleration g. Thus after a time t the body will have traveled a distance $AC = Vt$ in the direction AB, but it will also have fallen from C to D a distance $s = \frac{1}{2}gt^2$.
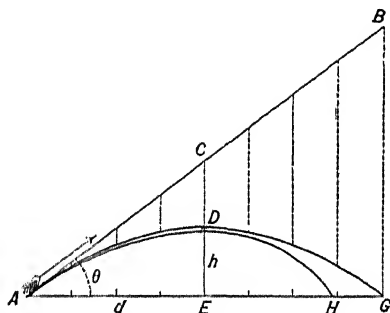
Fig. 53 Curve of projectiles and jets

If θ is the angle of elevation of AB above the horizontal, and if d is the distance AE which the projectile has advanced in a horizontal direction and h is its height, we have

$$d = Vt \cos \theta$$
$$h = Vt \sin \theta - \frac{1}{2}gt^2.$$

The path traversed may be shown to be a parabola with its axis vertical and passing through the highest point of the path. The highest point is halfway between the point of projection and the point G where the projectile again reaches the earth. The distance AG is called the range, and is a maximum when the angle θ is 45°.

These results are easily deduced from the above equations, but it must be borne in mind that the influence of air resistance has been neglected. This force in rapidly moving bodies, like bullets, may be very great and

changes the form of the trajectory to something like that shown by the second curve from A to H. In consequence of this the maximum range in gunnery is found at a much smaller elevation.

The form of the path of a projectile or ball is beautifully shown by a water jet, for each particle in the jet is a freely falling body.

112. Curved Pitching. If a ball when thrown forward is rapidly rotated the resistance of the air causes it to swerve from the path that it would otherwise take. This is seen in the curving of a pitched ball and in the drifting of projectiles from rifled guns. It results from the viscosity of air in consequence of which the rotating ball drags air in on one side and flings it out on the other as it advances.
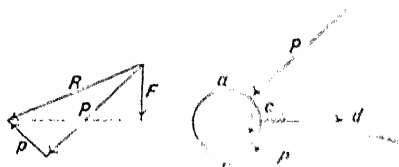

FIG. 54 Curving of pitched ball

Suppose, for example, that a ball is rotating about an axis perpendicular to the paper as shown by the curved arrow in figure 54, while it is moving forward in the direction of the straight arrow cd; the rotation of the ball drags air in from the side at a and carries it around toward the front of the ball at c, giving it a greater forward momentum than is given to the air between c and b where the surface of the ball is moving backward.

The force against the ball is, therefore, greater between a and c than it is between b and c. Let P and p represent these pressures against the ball. Their resultant as shown in the diagram of forces is the oblique force R which in part resists the forward motion of the ball, but also has a component represented by F which is at right angles to the path of the ball and causes it to swerve to one side in the direction of the dotted line.

As the force F acts *constantly* it causes the ball to move sidewise with constantly *accelerated* motion, and, therefore, the curving rapidly increases as the ball advances.

113. Motion Around the Earth. Suppose it were possible to shoot a cannon ball in a horizontal direction from the top of some high mountain on the earth with a velocity so great that while it advanced a mile it would drop just enough to follow the curvature of the earth. Then, if there were no air resistance, the

ball would continue around the earth and return to its original point of projection with undiminished velocity and would, therefore, continue to circulate forever around the earth as a satellite.

For suppose A (Fig. 55) is the point from which the projectile is shot in the direction AB. As it advances it drops away from the line AB; but the earth's surface also drops away from AB in consequence of its curvature, by about 8 in., in the first mile. If, therefore, the cannon ball has a velocity which will carry it a mile in the same time that it will drop 8 in., it will, on reaching the end of the mile, say at C, be just as high above the earth as
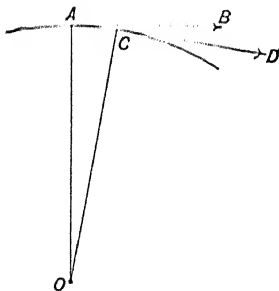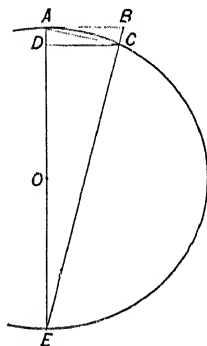
FIG. 55   FIG. 56

at the start and be moving in the direction CD, tangent at C. As the ball moves forward the force due to the earth's attraction is always at right angles to the direction of motion, and hence the speed of the ball is neither increased nor diminished and all the conditions of the motion remain constant.

The time required for a body to drop 8 in. toward the earth is found from the formula (§ 101)

$$s = \tfrac{1}{2}gt^2.$$

Since $s = 8$ in., or 0.66 ft., and taking $g = 32$ ft./sec.^{2}, we find $t = 0.20$ sec.; hence our cannon ball must have a velocity of a mile in 0.2 sec. or 5 miles per second.

The complete calculation may be made thus. Let the ball drop a distance $BC = s$ (Fig. 56) in going forward the distance $AB = d$. If R is the radius of the earth the relation between

s and d may be found from the similarity of the triangles ACE and ADC from which we find

$$AD : AC :: AC : AE$$

or

$$s : d :: d : 2R$$

and

$$\frac{d^2}{2R}$$

but s is the distance fallen with constant acceleration g in seconds, therefore

$$s \quad \tfrac{1}{2}gt^2,$$

and d is the distance which the ball, moving with constant velocity v, advances in t seconds; that is $d \quad vt$.

Substituting in (1) we have

$$\tfrac{1}{2}gt^2 \quad \frac{v^2t^2}{2R}$$

whence

$$v^2 \quad gR$$

Taking $g = 32$ ft./sec.^{2} and $R = 5280 \times 4000$ ft.,

$$v^2 = 32 \times 5280 \times 4000$$

and we obtain

$$v = 26{,}000; \text{ ft./sec.} = 4.92 \text{ miles per sec.}$$

114. Motion in Any Circle with Constant Speed. The case just discussed is in no way different from *any case when a mass moves in a circle with constant speed.* To cause it to change its direction of motion *there must be a force acting at right angles to the direction of motion, and if the speed does not change there can be no force at all in the direction of motion.* The relation between the velocity in the circle, the radius of curvature of the path, and the acceleration are given as above in the formula $a = \dfrac{v^2}{R}$.

It will be interesting to derive this relation in another way, from the simple conception of acceleration as the change in velocity per second.

As a particle moves from A to B in the circle (Fig. 57) its velocity changes only in *direction* from v_1 to v_2. But this change in velocity is equivalent to compounding with the original velocity v_1, another velocity represented by the vector f. This vector therefore represents the change in velocity between A and B and when divided by t, the time taken by the body in moving from A to B, gives the *average rate of acceleration between* A and B or $a : \dfrac{f}{t}$.
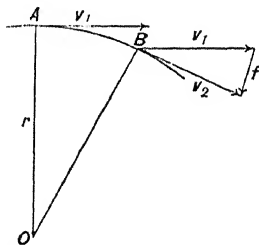
FIG. 57

Let d represent the distance AB, and since the sector OAB is almost exactly similar to the triangle formed by v_1, v_2, and f we have the proportion $r : d : : v : f$ where v is the *amount* of v_1 or v_2. But the distance $d = vt$ and $a = \dfrac{f}{t}$ or $f = at$, hence substituting

$$ r : vt : : v : at $$

from which

$$ a = \frac{v^2}{r}; $$

and this relation is *exact* and not approximate, for as B approaches A, the triangle and sector approach exact similarity as a limit and the *average* acceleration between A and B approaches the actual acceleration at A. It will be noticed also that f is parallel to a line bisecting the angle AOB, hence as B approaches A the direction of f approaches the direction AO as a limit. We conclude therefore that *the acceleration at any point of the circle is directed toward the center and is equal to* $\dfrac{v^2}{r}$.

115. Acceleration in any Curved Path. Since the acceleration just found depends only on the instantaneous relation of the various quantities involved, it applies to any curved path whatever. The acceleration at any point in a curved path may

be resolved into two components, one along the curve or tangent to it and the other at right angles to it. The component *along* the curve is the rate of change of speed of the moving body,
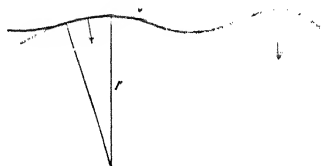
while the component at right angles to the path, i............ where

FIG. 58

v is the speed at the given point and r is the radius of curvature of the path at that point.

116. Force in Circular Motion. Whenever a mass is accelerated it is acted on by a force defined by the relation $F = ma$, hence when a mass m moves in a circle with constant velocity it is acted on by a force $F = m \dfrac{v^2}{r}$ directed toward the center of the circle; or, more generally, whenever a mass is moving in a curved path it is subject at any point to a force

$$F = m \frac{v^2}{r}$$

directed toward the center of curvature, r being the radius of curvature of the path and v the velocity of the mass m at the given point.

117. Centripetal and Centrifugal Force. The force by which a mass is constrained to move in a curved path is, as has just been shown, always directed toward the center of curvature of the path, and is therefore called the *centripetal* force. The reaction against this force by the moving body is called *centrifugal* force. Both are different aspects of the same stress, and are of course equal and opposite. For example, when a weight is whirled in a circle by a cord, it is held in the circle by the tension of the cord which supplies the *centripetal* force, while the reaction or outward pull of the weight against the cord is called the *centrifugal* force.

FIG. 59

When the string breaks *both* forces instantly disappear, *there is no tendency for the weight to fly outward*, it simply keeps moving in the tangential direction in which it was moving when freed.

118. Other Expressions for Centripetal Force. If a mass moves in a circle and makes n complete revolutions per sec-

ond, since the distance traveled in one revolution is $2\pi r$, in n
revolutions it will be $2\pi rn$ and hence $v = 2\pi rn$, and the centri-
petal force $F = \dfrac{mv^2}{r}$ becomes

$$F = m4\pi^2 n^2 r. \tag{1}$$

Or we may express the velocity in terms of the time required to
make one complete revolution, which may be called the *period*
T. In that case

$$v = \frac{2\pi r}{T}$$

and

$$F = m\,\frac{4\pi^2 r}{T^2}. \tag{2}$$

Or if ω represents the angular velocity of the rotating body, or
the arc in radians traversed per second, we have $\omega r = v$ (§ 136)
and therefore

$$F = m\omega^2 r. \tag{3}$$

119. Illustrations. When a railway train rounds a curve it
is kept in the curve by the pressure of the rail against the flanges
of the wheels. The weight of the train
is balanced by the upward pressure of
the track, represented at A, figure 60,
while the centripetal force exerted by
the rails is represented by B; the re-
sultant force R is therefore inclined and
the track is tilted toward the center so
that the pressure may be equal on both
rails.
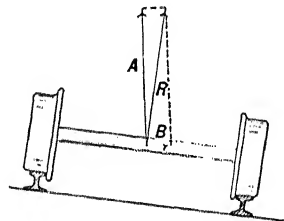
FIG. 60

When a fly wheel rotates, it is under
great tension due to the centrifugal force of the heavy rim, and
great destruction may result from the bursting of such a wheel.
A grindstone also may burst if driven at too high a speed.

A glass of water held in a sling may be swung in a vertical
circle without spilling the water. For the acceleration down-
ward due to gravity when it is at the top of the circle may not be
enough to hold it in the circular path, consequently the bottom
of the glass must exert an additional force upon the enclosed

water, so that even at the top of its path the water presses
against the bottom of the glass. In this case the tension on the
cord when the mass is at the top of the circle is lessened by the
weight of the body, while at the bottom the tension is increased
by the same amount. The tension on the cord will therefore be

$$F = \frac{mv^2}{r} - mg \qquad \text{at the top.}$$

$$F = \frac{mv^2}{r} + mg \qquad \text{at the bottom.}$$

120. Conical Pendulum. Suppose a mass m, as in an old-
fashioned steam engine governor, is swung around in a circle
with uniform speed, it will swing out from the axis and come into
equilibrium at a certain angle depending on the speed of rotation.
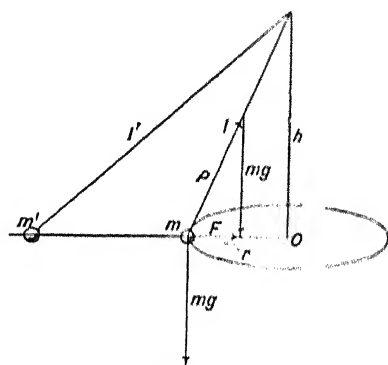
Evidently in order that
there may be equilibrium the
mass must be acted on by a
force F directed toward the
axis and just sufficient to hold
it in the circle. But if the
mass, or conical pendulum as
it may be called, makes one
revolution in T seconds, then
the centripetal force is

$$F = \frac{m 4 \pi^2 r}{T^2},$$

FIG. 61

and this force is the resultant
of the tension on the cord P and the weight of the mass, which
is mg dynes. Constructing the diagram of forces as in the figure,
it is clear that $F : mg : : r : h$; therefore,

$$F = \frac{mgr}{h}.$$

We have, therefore,

$$\frac{mgr}{h} = \frac{m 4 \pi^2 r}{T^2},$$

whence

$$T = 2\pi \sqrt{\frac{h}{g}}.$$

Consequently if we have two masses m and m' hung by cords of different lengths, they will have the same period of rotation if the height h is the same in both cases.

The more rapid the rotation the higher the mass rises, and in the steam-engine governor the rising of the weights operates a lever through which steam is cut off and the speed of the engine decreased.

PROBLEMS

1. A stream of water from a horizontal nozzle falls 3 ft. below the level of the nozzle in a distance of 20 ft., measured horizontally. Find the velocity of the escaping jet.

2. A jet of water is directed upward at an angle of 45° to the vertical, and strikes the ground at a distance of 64 ft. from the nozzle. Find the time taken by a water particle in passing from nozzle to ground, and the velocity of the jet.

3. How much weight can a cord sustain by which a mass of 100 gms. can be whirled in a circle of 1 meter radius making 2 turns per sec. neglecting the effect of gravity in the circular motion?

4. A stone weighing 1 lb. is whirled by a string in a circle 6 ft. in diameter. The string breaks and the stone flies off with a velocity of 30 ft. per sec. Find the strain on the string when it broke.

5. A mass of 50 gms. has a velocity of 750 cms. per sec. in a circle of radius 60 cms. Find the acceleration in amount and direction and centripetal force in dynes. Also find angular velocity.

6. A 100-ton locomotive rounds a curve at a uniform speed of 40 miles per hour. Find the acceleration if the radius of curvature of the track is 1000 ft. Also find the horizontal force exerted against the rails.

7. In case of the last problem, how much higher must the outer rail be than the inner, in order that the resultant force, due both to the weight of the locomotive and its centrifugal force, may be perpendicular to the road bed?

8. A mass of 1 lb. is whirled in a circle of 2 ft. radius on a smooth level table, being held in the circle by a cord which passes without friction through a hole in the center of the table and supports a 2-lb. weight. Find the angular velocity and revolutions per sec. of the 1-lb. mass necessary to support the weight.

9. A 200-gram mass is whirled in a vertical circle of radius 80 cms. with a uniform angular velocity 8 radians per sec. Find the period of revolution and the acceleration. Also what is the tension on the cord in grams when the mass is at the top of the circle and when it is at the bottom?

10. A weight of 2 lbs. is whirled in a vertical circle. If its velocity is 100 cms. per sec. at the top of the circle, what will be its velocity at the bottom, the gain being due to the acceleration of gravity as it falls, as if as in an inclined plane (see § 106)? Radius 80 cms.

11. A 10-lb. mass is hung as a pendulum by a cord 4 ft. long. How high must it swing in order that the tension on the cord at the lowest point of its swing may be double the tension when hanging at rest.

12. In case of "looping the loop," how high above the head of the top of the circle must the car start that it may just have speed enough to keep in the circle, neglecting friction? Circle 30 ft. in diameter.

13. Find the angular velocity and period of a conical pendulum hung by a cord 1 meter long and swinging around in a horizontal circle of 60 cms. radius.

PERIODIC AND VIBRATORY MOTION

121. Introduction. Any motion which exactly repeats itself through equal intervals of time is called a *periodic* motion. Many of the most familiar types of motion are periodic, including the motions of the planets in their orbits around the sun, the vibration of a weight on a spring, or the motion of the prongs of a tuning fork. The swing of the pendulum of a clock or the oscillations of a balance wheel of a watch will also be recognized as periodic motions. All wave motions are the result of periodic motions of individual particles of the medium which transmits the waves. If a periodic motion, *when once started, repeats itself without the aid of any external agency,* usually dying out only gradually, it is called a *vibration.* In the sections which immediately follow, vibratory motion will be discussed.

122. Condition Necessary for a Vibration. In order that a given mass or a particle of material shall vibrate, a displacement from its *rest position* must be opposed by a so-called *restoring force.* When displaced and let go, the body is continually urged back toward its rest position by this restoring force. The kinetic energy acquired during its motion carries the body beyond the rest position and the body may perform a very large number of vibrations before finally coming to rest. A pendulum bob will swing hundreds of times before its vibration ceases. If the vibration is *linear,* that is to say, if it is performed in a straight line, it passes through its rest position twice during every complete vibration cycle, once in its forward swing

and once in its backward swing. If the vibration is circular or elliptical the vibrating body never passes through its rest position but continually swings around it, reaching it only when it finally comes to rest.

The vibration may be purely rotational in character as in the case of the balance wheel of a clock or watch. The restoring force is in this case a turning force which tends to rotate the balance wheel about its axis toward its rest position. Like a linear vibration a body performing a rotational vibration swings through its rest position twice during every vibration cycle.

If no friction of any kind is present to resist the motion of the vibrating body, its vibration, when once started, will continue indefinitely without further stimulus. Most vibrations with which we are familiar, such as a swinging pendulum, or a weight vibrating up and down on the end of a spring, are observed to cease gradually, that of the pendulum chiefly through air resistance, and that of the weight on the spring through air resistance and an internal friction within the metal of the spring itself. The case of the planets revolving about the sun is one in which the frictional forces are relatively so small that the periodic motion continues indefinitely.

In every vibration (§ 121) *a restoring force is present tending to restore the vibrating body to its position of rest.*

123. The Restoring Force. The restoring force which produces the simplest and the commonest types of vibration is one which is directly proportional to the distance which the body is displaced from its rest position. If a weight suspended upon the end of a helical spring is lifted above or pulled below its rest position, measurement will show that the amount of force required is directly proportional to the displacement produced, whether the displacement be above or below the rest position. The same is true of sideways displacements of the pendulum bob for small displacement angles (§ 131), and of the angular displacements of the watch balance wheel; in both of these cases the restoring force is directly proportional to the displacement.

When the vibration produced by a restoring force of this character is *linear* or very nearly so as in the case of the pendulum bob the motion is said to be *simple harmonic.* Simple harmonic

motion will be defined and discussed in some detail in the following paragraphs.

124. Simple Harmonic Motion. If a ... moves ... th constant speed in a circular path is observed ... oint in the plane of the circle, it appears to ... ath in a straight line.

The kind of vibratory or oscillatory motion the particle appears to have in this case is known as ... motion; it may ... tion upon a stra... form motion in a ...
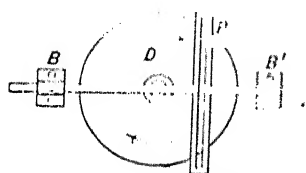

FIG. 62

There are other kinds of vibratory motion that are *not* simple harmonic, such, for example, as the particle would appear to have in the above instance if it moved around the circle in any manner whatever *except* with constant speed. Simple harmonic motion is, therefore, one particular mode of oscillation; but it is by far the most important, for it is the most common of all, and all other modes of vibratory motion may be expressed as the resultant of a sum of simple harmonic motions, as was shown by the French mathematician Fourier.

A simple mechanical device illustrating this kind of motion is shown in figure 62. A pin P projects from the face of a rotating disc D and fits in a slot in a cross head which is attached to rods that can slide back and forth in the bearings BB'. When the disc rotates with uniform speed every point in the rods and cross head will move back and forth with simple harmonic motion.

FIG. 63

The *amplitude* of a vibration is the distance that the vibrating body moves on each side away from its central or mean position.

125. Velocity in Simple Harmonic Motion. Let a particle A move around the circle (Fig. 63) with constant speed, and let another particle B move back and forth along a diameter DC
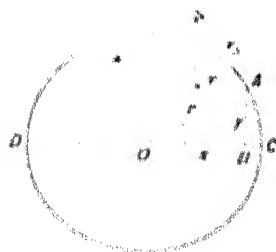
in such a way that the line joining A and B is always perpendicular to DC. Then B oscillates with simple harmonic motion. Let v_0 represent the velocity of A. It may be resolved into two components, as shown in the diagram, one at right angles to the direction in which B moves and the other parallel with B's motion. Since B always keeps abreast of A, the velocity of B at any point must be equal to that component of A's velocity which is parallel to DC, namely to the component v. Letting r represent the radius of the circle and y the distance AB, we have by similar triangles

$$r : y = v_0 : v$$

whence

$$v = \frac{\cdot}{r} v_0 \quad \text{or} \quad v = v_0 \, sin \, e$$

where e is the angle AOC.

The velocity of B is, therefore, zero at the ends of its path at C and D, for there $y = 0$. While at the center $y = r$ and the velocity of B is equal to v_0, its maximum value.

The complete period of an oscillation of B is evidently the same as the time in which A goes completely around the circle. Let T represent this period, and the velocity of A is

$$v_0 = \frac{2\pi r}{T},$$

which also expresses the velocity of B at its middle point.

126. Acceleration in Simple Harmonic Motion. Since A and B have exactly the same motion in the direction DC, the acceleration of B must be the same as that component of the acceleration

FIG. 64

of A which is parallel to DC. The acceleration a_0 of A, moving with uniform speed in a circle, is directed toward the center of the circle and is equal to $\frac{4\pi^2 r}{T^2}$ (§ 118). Resolving the acceleration a_0 into two components and letting a represent the
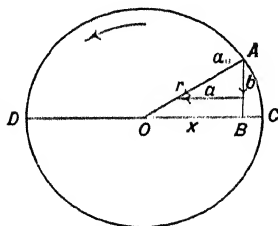
component parallel to DC and b that perpendicular to it, we have by similar triangles,

$$a : a_0 \quad x : r$$

and, therefore,

$$a \quad \frac{a_0}{r}\, x$$

or since

$$a_0 \quad \frac{4\pi^2}{T^2}\, r$$

$$a \quad \frac{4\pi^2}{T^2}\, x.$$

The acceleration of B is, therefore, proportional to its distance from the center, it is greatest when $x \quad r$ and is zero when B is at the center. It will also be noticed that the acceleration of B is always directed *toward* the center; that is, B is always losing velocity as it moves away from the center and gaining velocity as it moves toward the center, and consequently its velocity is greatest at the center as we have already seen.

127. Force in Simple Harmonic Motion. The fundamental dynamical equation $F \quad ma$ enables us to express at once the force in simple harmonic motion. When the mass of the oscillating particle is m and its complete period of oscillation is T, the acceleration at the instant when the particle is a distance x from its central position has just been shown to be

$$a \quad \frac{4\pi^2 x}{T^2}.$$

The force at that instant is, therefore,

$$F \quad m\frac{4\pi^2 x}{T^2}$$

and is directed always toward the center, or

Therefore when a mass in equilibrium is so situated that if displaced it is always urged back toward its equilibrium position by a force which is proportional to the displacement, it will, on being displaced and then set free, oscillate with simple harmonic motion about its position of equilibrium as a center.

Now the force required to cause a *small* strain in almost any elastic body is proportional to the amount that the body is strained, whether the body is bent or stretched or twisted (Hooke's Law, § 245), hence when such bodies are strained and then let go they oscillate to and fro in simple harmonic motion, as in case of the small vibrations of a tuning fork.

128. Problem. Let a mass of 1 kgm. be supported by a steel spring of such stiffness that an additional weight of 100 grams will stretch it just 1 cm. It is required to find the period of oscillation of the weight if disturbed, neglecting the mass of the spring.

If the kilogram weight is pushed up or pulled down as it hangs on the spring, it will move through a distance which is proportional to the force used, a force of 100 gms. being required to displace it 1 cm. To produce a displacement of x cms. the force required is $100x$ gms. or $100xg$ dynes. But from equation (1) above we have, since $m = 1000$,

$$F = 1000 \frac{4\pi^2 x}{T^2}$$

but

$$F = 100gx.$$

Therefore, $100g = 1000 \dfrac{4\pi^2}{T^2}$ and $T^2 = \dfrac{4000 \cdot \pi^2}{100 \cdot g}$

which gives $T = 0.634$ sec.

129. Simple Harmonic Motion Isochronous. It will be noticed that the expression $T^2 = \dfrac{4\pi^2 xm}{F}$ does not contain r and is, therefore, independent of the amplitude of the vibration, so that it does not make any difference in the period of vibration whether the amplitude is large or small, provided the ratio $\dfrac{x}{F}$ is constant, in which case the motion is truly simple harmonic.

When vibrations have this property they are said to be *isochronous*.

130. Energy of a Vibrating Mass. In every vibration the *sum of the potential and kinetic energies of the vibrating body remains constant, throughout the entire vibration cycle*, except during the building up or the damping out of a vibration when the total energy respectively increases or decreases. It can be seen that this must be true from the principle of the conservation of energy (§ 72) because when a body is vibrating freely, such as a swing-

132. Pendulum Clocks. The pendulum affords a valuable means of regulating the motion of a clock, since when it swings through a small arc its oscillations are nearly *isochronous*, i.e., its period of oscillation is nearly independent of the amplitude of its swing.

When an ordinary clock driven by a spring is just wound up it gives a greater impulse to the pendulum through the escapement than when it is nearly run down, and even in clocks driven by weights the friction is not always constant and the swing of the pendulum will vary accordingly.

It must be remembered also that the pendulum of a clock is not *free*, but the little backward and forward impulses which it receives from the escapement hasten somewhat its motion. To secure regularity of motion, therefore, the pendulum should be heavy, so that its natural period will be only slightly affected by the pushes of the escapement.

In good astronomical clocks what is known as a *gravity escapement* is used, in which the pendulum does not receive any impulse directly from the spring or weight that drives the clock, but its motion is kept up by a small weighted lever which is set free just as the pendulum reaches the end of its swing, and in falling gives a slight push to the latter.

Between successive impulses the lever is raised and set in position by the action of the clockwork.

Full details as to some forms of gravity escapement will be found in the article *Clocks* in the "Encyclopedia Britannica."

133. Free and Forced Vibrations. Before leaving the subject of vibratory motion the difference between a free and a forced vibration should be clearly understood. The term *free vibration*, which is often used, means, as its name implies, a vibration performed by a body which vibrates freely at its *natural period*. A forced vibration is an oscillation of a body maintained by an external periodic force in a period other than the natural period of the body. If the external force be removed, the body either ceases to oscillate or changes the period of its oscillation to that of its own natural period.

When the period of a vibration is referred to, it is always the natural period which is meant unless otherwise stated.

134. Resonance. The idea of *resonance* is well brought out in illustrating the difference between a forced and free vibration.

When a rubber ball suspended from an elastic cord is made to oscillate up and down by raising and lowering the top end of the cord slightly by the hand, it will be recalled that there is one particular period of up and down motion of the hand to which the ball actively responds. If the hand is oscillated more rapidly or more slowly than this period the response is very sluggish. The frequency for which the ball responds strongly is its free or natural period of up and down oscillation and the vibration is said to be in *resonance* with the stimulus of the hand. For all other frequencies of stimulus the ball fails to resonate and the feeble oscillations which are observed are forced vibrations. A resonant vibration builds up to a relatively large amplitude with a very small stimulus. Furthermore if the ball is in resonance with the motion of the hand, when the stimulus from the hand is removed by holding the hand stationary, the ball continues to oscillate freely up and down at this frequency for some moments. On the other hand the small up and down motion of the ball produced by the motion of the hand during the forced oscillations ceases as soon as the motion of the hand ceases and is replaced by a feeble temporary oscillation in the natural period of the ball. This is in accordance with the definition of a forced oscillation given in the preceding section.

Thus it is seen that resonance takes place at only one particular frequency, namely, the free vibration period, while forced oscillations of small amplitude may be produced at all other periods. Furthermore a resonant vibration may gradually build up to a large amplitude of considerable energy from a very small stimulus which supplies energy at a small rate.

Resonance is important in sound vibration phenomena (§ 318) and in electrical oscillations (§ 813).

The tuning of a radio circuit is simply the adjustment of the natural electrical vibration period of the circuit until it is in resonance with the incoming electric waves, so that a strong electrical oscillation is built up by the action of a small electrical stimulus. The oscillations produced in the circuit when it is out of tune are forced oscillations, so small in amplitude that they cannot be detected except in the proximity of a powerful transmitting station.

135. Vibration Formula. The formula for the period of vibration of a pendulum is a special case of a general formula

$$T = 2\pi \sqrt{\frac{m}{F_1}} \qquad (1)$$

which applies to all cases of a vibrating mass, acted upon by a restoring force which is proportional to the displacement of the mass from its rest position. In this formula

$T =$ period of the vibration in seconds

$m =$ mass of the vibrating body

$F_1 =$ restoring force per unit displacement of the mass from its rest position in absolute units.

For the case of the pendulum, using the previous notation $m = m$ and $F_1 = \dfrac{mg}{l}$ which gives the pendulum formula, on substitution in equation (1). The value of F_1 in this case is derived as follows: The downward force on the mass m equals mg. If the mass is displaced sideways unit distance, the restoring force as found by the diagram of forces equals $m g \cdot \dfrac{1}{l} = \dfrac{mg}{l}$. In this case, however, the sideways displacement must be small compared with the length of the pendulum as discussed in the small print of § 133. The period of an angular vibration such as that of the balance wheel of a clock is expressed by a formula exactly analogous to (1) as is also the period of an electrical oscillation (§ 149 and § 811).

PROBLEMS

1. Show that the motion of the piston of a steam engine when the crank is turning with uniform velocity is not simple harmonic. At which end of the piston's motion is the acceleration greatest and why?

2. Assuming that the motion of the piston is simple harmonic, find its velocity in the middle of its stroke when the crank is 8 in. long and makes 200 revolutions per minute. Also find acceleration at middle and end of its stroke.

3. If the piston and connecting rod weigh 100 lbs. in the last problem, find the maximum force against the crank pin due to their inertia alone, neglecting the effect of steam pressure.

4. A mass of 4 lbs. is made to oscillate to and fro by a spring at the rate of 2 vibrations per sec. Find the force on the mass when it is 2 in. from its middle position.

5. A pendulum 1 meter long swings 10 cms. on each side of its lowest point; find the direction and amount of the acceleration at the ends of its swing and at middle.

6. How long must a pendulum be to beat seconds at a place where $g = 980$? If made 1 mm. too long will it gain or lose and how much per day?

7. A clock having a pendulum which beats seconds where $g = 980$, is taken to another place where $g = 981$; will it gain or lose, and how much in one day?

8. Each prong of a tuning fork, making 100 complete vibrations per second, vibrates to and fro through a distance of 1.5 mm. Find the velocity of the prong in the middle of its swing.

9. A 400-gm. weight when hung on a long and light helical spring stretches it 30 cms. What will be its period of oscillation if drawn down a little and then set free? Take $g = 980$ and neglect mass of spring.

Ans. 1.099 sec.

IV. ROTATION OF RIGID BODIES

Motion of a Rigid Body

136. Translation and Rotation. If a rigid body moves in such a way that any straight line joining two points in the body remains parallel to itself as it moves along, the motion is said to be a *translation* without rotation. A book slid about on a table with one edge always parallel to one edge of the table is a case of pure translation. If the edge of the book changes its direction there is said to be *rotation*.

Any motion of a rigid body may be considered as made up of the motion of its center of mass combined with rotation about an axis through that center.

Motion of the Center of Mass. It may be proved that when any *external* forces act on a rigid body the center of mass of the body moves just as though the whole mass of the body were concentrated at that point and all the forces were applied directly to it, and it makes no difference at what points on the body the forces may be applied.

When a top spins on a smooth frictionless table its center of gravity remains at rest, for the external forces acting on the top are its *weight* due to the attraction between it and the earth and the upward *pressure of the table on its point*. These two forces are equal and opposite and consequently the center of gravity has no translational acceleration even when the top is inclined as in figure 75 (p. 109).

When a stick of wood is hurled through the air its center of mass moves in a simple parabolic curve (§ 111) just as a particle would move, except as affected by air resistance.

Besides this translational force which depends only upon the amounts and directions of the several forces and to is found by the simple force polygon, *there is usually also a couple which causes the body to rotate about an axis through its center of mass.* This couple depends not only on the amounts and directions of the forces, but also upon their points of application to the body.

137. Angular Velocity. When any line in a body is at rest while other points in the body move in circles about that fixed line or *axis*, the motion is called *rotation.* In case of a rigid body, like a wheel, all parts whether near the axis or far from it must rotate through equal angles in the same time. The rate at which the body is turning at any instant, measured in *radians per second*, is known as its *angular velocity* and is represented by ω (the Greek letter *omega*).

Since the length of a radian of arc at a distance r from the axis is equal to r, we have

$$\omega r$$

where v is the linear velocity of a particle at a distance r from the axis.

EXAMPLE: If a wheel of radius 15 cms. is making 3 revolutions per sec., its angular velocity is $3 \times 2\pi$ radians per sec., and the linear velocity of a point on the rim is $3 \times 2\pi \times 15$ cms. per sec.

138. Angular Acceleration. When the angular velocity of a body is changing, the rate of change per second is known as its *angular acceleration*, and may be represented by α. If the angular velocity ω_1 changes to ω_2 in t seconds, then

$$\alpha = \frac{\omega_2 - \omega_1}{t}$$

or

$$\omega_2 = \omega_1 + \alpha t$$

where α is the average rate of acceleration during the time t.

The *direction* of the axis of rotation may change, and this also constitutes an *angular acceleration*, even though the speed of rotation about the axis may remain constant. This is illustrated by the motion of a spinning top when its axis is inclined, for the axis swings around in a circle keeping a constant inclination to the vertical.

139. Vector Representation of Angular Velocity. The angular velocity of a body may be represented by a vector or arrow drawn along the axis of rotation and having a length proportional to the amount of the angular velocity, and pointing in the direction that a person must look along the axis to see the body rotating in a clockwise direction. For example, if the rotating disc shown in figure 66 has an angular velocity 10, it will be represented by a vector 10 units long drawn in the direction of the arrow.
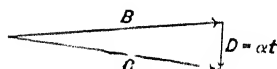
140. Change in Direction of Angular Velocity. If the axis of rotation is changing in direction the angular velocity at one instant might be represented by the vector B and a short time

FIG. 66

FIG. 67

later by C (Fig. 67). The change in angular velocity would then be represented by the vector D, for this combined with B gives C according to the composition of vectors. If t is the time during which the change has taken place, then $D = \alpha t$ where α is the *angular acceleration*. An angular acceleration of this character is found in the motion of a top (§ 155).

141. Rotation with Constant Acceleration. The equations for rotation with constant acceleration are exactly analogous to those for simple translation (§ 98), as may be seen thus:

Translation in Straight Line	*Rotation about a Fixed Axis*
$s =$ displacement in time t.	$\theta =$ angle through which body turns in time t.
$v_1 =$ velocity at *beginning of* interval t.	$\omega_1 =$ angular velocity at *beginning* of interval t.
$v_2 =$ velocity at *end* of interval t.	$\omega_2 =$ angular velocity at *end* of interval t.
$a =$ acceleration.	$\alpha =$ angular acceleration.

$$a = \frac{v_2 - v_1}{t}. \qquad\qquad \alpha = \frac{\omega_2 - \omega_1}{t}. \tag{1}$$

$$\left(\frac{v_2 + v_1}{2}\right) t \text{ or } s = v_1 t + \frac{at^2}{2}. \qquad \theta = \frac{\omega_2 + \omega_1}{2} t \text{ or } \theta = \omega_1 t + \frac{\alpha}{2} t^2. \tag{2}$$

$$2as = v_2^2 - v_1^2. \qquad\qquad 2\alpha\theta = \omega_2^2 - \omega_1^2. \tag{3}$$

PROBLEMS

1. If a wheel revolves 1800 times per minute, what is its angular velocity; and if it is 6 in. in diameter what is the linear velocity of a point on its periphery?

2. What is the linear velocity of a point 1 ft. from the axis of a wheel making 2.5 turns per sec.? Also the velocity of a point 1 4 ft. from axis? What is the angular velocity of each?

3. Find angular velocity of a wheel in which a point 6 in. from the axis has a velocity of 4 ft. per sec.

4. A locomotive rounds a curve having a radius of 800 ft. at 15 miles per hour; what is its angular velocity?

5. A wheel is given a speed of 100 revolutions per min. in 2 minutes; what is its angular acceleration in radians per sec. per sec.

6. How many revolutions will a fly wheel make in 20 seconds, while its angular velocity is changing from 3 to 10 radians per sec., if the acceleration is constant?

7. A body rotates about an axis with constant angular acceleration 8 radians per sec. per sec.; how many turns will it have made in 10 seconds from the start?

8. How many revolutions will a body make starting from rest with angular acceleration 4 radians per sec. per sec. before it will be revolving at the rate of 20 turns per sec.?

KINETICS OF ROTATION ABOUT A FIXED AXIS.

142. Angular Acceleration Caused by Torque. Suppose the bar shown in figure 68 is acted on by a force F at a distance d from the axis; it is required to find how rapidly the speed of rotation of the bar about the axis will increase in terms of the moment of force, or torque Fd.

Imagine the bar divided into little masses m_1, m_2, m_3, etc., and suppose the effect of the force F is to cause an *angular acceleration* α in the rotation of the bar; that is, its angular velocity is increased at the rate of α radians per sec. per sec. The *linear* acceleration of the mass m_1 at distance r_1 from the axis will then be $r_1\alpha$, and consequently the force acting on m_1 must be $m_1 r_1 \alpha$ and may be represented by f_1. This force f_1 is due to F and is transmitted to m_1 by the rigidity of the bar. So also m_2 must be acted on by a force $f_2 = m_2 r_2 \alpha$ since it has the acceleration $r_2\alpha$.

And similarly every one of the masses m_1, m_2, m_3, etc., into which the bar is divided is acted on by the force needed to give it its acceleration, as indicated by the small arrows in the figure.

Now, if a force equal and opposite to f_1 is applied to m_1, and a force equal and opposite to f_2 is applied to m_2, and so on, applying to each of the little masses a force just such as to counteract its acceleration, it is clear that there will be no acceleration and the bar will be in equilibrium. That is, a system of forces equal and opposite to

FIG. 68

f_1, f_2, etc., will just balance the turning moment of the force F about the axis O. Consequently the sum of the moments of f_1, f_2, etc., about O must be equal to the moment of F about that axis. Thus,
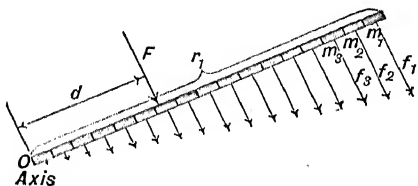
$$Fd = f_1 r_1 + f_2 r_2 + f_3 r_3 +, \text{ etc.}$$

But it has been shown that

$$f_1 = m_1 r_1 \alpha, \quad f_2 = m_2 r_2 \alpha, \text{ etc.}$$

Therefore

$$Fd = m_1 r_1^2 \alpha + m_2 r_2^2 \alpha + m_3 r_3^2 \alpha +, \text{ etc.,}$$

or

$$Fd = \alpha \left(m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 +, \text{ etc.} \right).$$

The quantity in the parenthesis, which depends only on the mass of the body and its distribution with reference to the given axis, is called the *moment of inertia* of the body about that axis and may be represented by the symbol I.

The *torque* or sum of the moments of whatever forces may be acting to rotate the body around the given axis may be represented by L, and we have then,

$$L = I\alpha \quad \text{or} \quad \alpha = \frac{L}{I}. \tag{1}$$

That is, *the angular acceleration caused by a given torque is equal to the torque divided by the moments of inertia of the body about the given axis.*

Notice the analogy to the formula F ... of ... torque corresponds to force, moment of inertia corresp... angular acceleration corresponds to linear acceleration.

The effect of torque in causing angular acceleration may be illustrated by the apparatus shown in the figure. A light bar carrying two masses M and M' is mounted on a horizontal axis perpendicular to the bar and is set in motion by a ... light H hung from a cord wrapped ar... ... on the axis. When the masses M and M' are in the position shown, the bar gains angular velocity slowly, for the farther the ma... es are from the axis, the greater the moment of inertia of the rotating system. When the masses are close to the axis the moment of inertia is smaller and the bar gains angular velocity very much more rapidly than before.
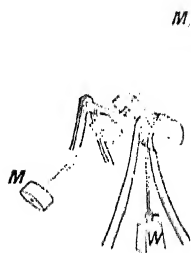
The calculation of moments of inertia will be discussed in paragraph. 145 to 147.

143. Angular Momentum.

The formula of the last ...

$$L. = I\alpha$$

may be put in the form

$$L. = I^{\omega_2} \qquad \text{§ 14})$$

or

$$L. = I\omega_2 = I\omega_1 \qquad (1)$$

which is exactly analogous to

$$Ft = mv_2 = mv_1 \qquad (\text{§ 94})$$

The product of the moment of inertia by the angular velocity about an axis is known as the *angular momentum* of the rotating body about that axis, and equation (1) above, states that the *change in the angular momentum of a body about any axis is equal to the moment of force or torque about that axis multiplied by the time during which it acts.*

When the axis of torque is *perpendicular to the axis of rotation* of the body its only effect is to change the *direction* of the axis of rotation, but the *amount* of the angular momentum remains unchanged. This is illustrated by the top (§ 155).

FIG. 69

144. Kinetic Energy of a Rotating Body. When all parts of a body have the same velocity the kinetic energy of the body as we have already seen is $\frac{1}{2}Mv^2$ where M is the mass of the body and v its velocity. But in case of a rotating rigid body the velocity of any part depends on its distance from the axis. In this case we may imagine the whole mass to be divided into small portions, and calculate the kinetic energy of each of these portions separately and then add them together to find the total energy of rotation.

The body represented in figure 70 is supposed to rotate about an axis perpendicular to the paper. Imagine the whole body cut up into little rods parallel to the axis whose ends are seen as the reticulation in the diagram. Let the mass of one of these rods be m, its distance from the axis r, and its velocity due to the rotation of the body v. Then its kinetic energy is $\frac{1}{2}mv^2$.
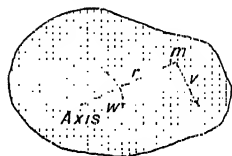

Fig. 70

But if ω is the *angular velocity* of the body, ωr will be the linear velocity of a mass at a distance r from the axis.

Thus,

$$\omega r = v \quad \text{and} \quad \tfrac{1}{2}mv^2 = \tfrac{1}{2}m\omega^2 r^2.$$

Now, let m_1 represent the mass of another of the rods into which the body has been imagined divided and r_1 its distance from the axis, then its kinetic energy is $\frac{1}{2}m_1\omega^2 r_1^2$, and so the total kinetic energy of the body is

$$K. E. = \tfrac{1}{2}m\omega^2 r^2 + \tfrac{1}{2}m_1\omega^2 r_1^2 +, \text{ etc.,}$$

there being one term for each part into which the body is conceived to be divided. Or we may write

$$K. E. = \tfrac{1}{2}\omega^2(mr^2 + m_1 r_1^2 +, \text{ etc.}),$$

since the angular velocity of every part of the body is the same. But the quantity in parenthesis is the moment of inertia I of the bar about the axis, therefore

$$K. E. \quad \tfrac{1}{2}I\omega^2.$$

Notice again the analogy between this expre   and the formula for kinetic energy of translation $\frac{1}{2}MV^2$.

Moment of inertia corresponds to mass
Angular velocity corresponds to linear velocity

145. Moment of Inertia of a Rod. The method of com   ments of inertia may be illustrated by the case of a straight uni   th the axis at one end. Let l be the length and M the mass of the   ceive it to be divided into n equal parts, each part having a ma   on the length of each part will be $\frac{l}{n}$, and if the distance of any pa   is taken as the distance of its farther end, the dista   parts are $\frac{l}{n}$ $\frac{2l}{n}$ $\frac{3l}{n}$, etc., and the moment of inertia is, thereb

$$I = m\frac{l^2}{n^2}\quad m\frac{2^2l^2}{n^2}\quad m\frac{3^2l^2}{n^2}\ +, \text{ etc., or } I\quad \frac{ml^2}{n^2}\quad n^2).$$

Now, it may be shown that the larger n is taken the   re closely does the sum in the parenthesis approach the value $\frac{n^3}{3}$, and as   dingly if the rod is supposed to be divided into an infinite number of parts,

$$I\quad \frac{ml^2}{n^2}\cdot\frac{n^3}{3}\quad \frac{Ml^2}{3}\ \text{ since } mn = M$$

The moment of inertia of the bar is, therefore, the same as though its mass were concentrated at a distance k from the axis, where $k^2 = \frac{l^2}{3}$

The distance k is known as the *radius of gyration* of the rod about the given axis.

146. Formulas for Moment of Inertia. In case of bodies of simple figure and having the mass uniformly distributed throughout the volume the moments of inertia may be calculated by the methods of calculus. But in more complicated cases they must be determined by experiment.

The following formulas are given for reference:

Thin rod, of mass M and length l, having a transverse axis at one end,

$$I\quad \frac{Ml^2}{3}.$$

Thin rod, of length l, having a transverse axis through the center,

$$I = \frac{Ml^2}{12}.$$

Rectangular block, of width a and length b and of any thickness whatever, about an axis through the center perpendicular to a and b,

$$I = M \left(\frac{a^2 + b^2}{12} \right)$$

Circular disc or cylinder, of any length and of radius r, about an axis through the center and perpendicular to the circular section of the disc or cylinder,

$$I = \frac{Mr^2}{2}.$$

Circular cylinder, of length l and radius r, about a transverse axis through its center perpendicular to its length,

$$I = M \left(\frac{r^2}{4} + \frac{l^2}{12} \right).$$

Sphere, of radius r about an axis through its center,

$$I = M \frac{2r^2}{5}.$$

147. Moment of Inertia about a Parallel Axis. If the moment of inertia of a body is known about an axis through its center of mass, it may readily be calculated about any parallel axis. For if I_0 is the moment of inertia about the axis through its center of mass and if M is the mass of the body, then the moment of inertia about a parallel axis at a distance h from the center of mass of the body is $I = I_0 + Mh^2$; that is, the moment of inertia I about any axis is equal to the moment of inertia which the whole mass would have about that axis if it were concentrated at the center of mass of the body, added to the moment of inertia of the body about the parallel axis through its center of mass.

FIG. 71

148. The Compound Pendulum. In discussing the *simple pendulum* it was assumed that the oscillating mass was so small that it might be considered as concentrated at a point, and the mass of the suspending system was entirely neglected.

A pendulum which has distributed mass and so does not satisfy either of the above simple conditions is said to be a *compound* or *physical pendulum*. All actual pendulums belong to this class.

Let it be required to find the length of a simple pendulum having the same period of oscillation as a given physical pendulum. Suppose the pendulum to have mass M and let its axis of suspension O be a distance h above its center of gravity C (Fig. 71). Then, when a line joining O and C makes an angle θ with the vertical, the pendulum may be considered as acted upon
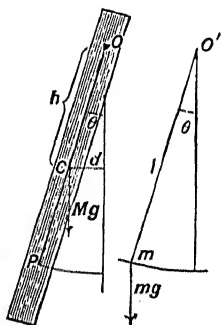
by a force Mg acting downward through its center of gravity and producing a moment of force about the axis O equal to $M \dots$ $M \dots$. If I is the moment of inertia of the pendulum about O, we have by equation 1 §142,

$$Mgh \sin \theta = I\alpha,$$

therefore,

$$\alpha = \frac{Mgh \sin \theta}{I}$$

But in case of a simple pendulum of length l the moment of the force mg about the axis O' is $mgl \sin \theta$ and the moment of inertia of m about O' is ml^2; therefore, $mgl \sin \theta = ml^2\alpha'$ and the angular acceleration is

$$\alpha' = \frac{g \sin \theta}{l}.$$

If the two pendulums are to have the same period of vibration their angular accelerations α and α' must be equal when both pendulums make equal angles with the vertical; that is,

$$\frac{Mgh \sin \theta}{I} = \frac{g}{l} \sin \theta$$

and, therefore,

$$l = \frac{I}{Mh}$$

The length of the equivalent simple pendulum calculated from the above formula will always be greater than h, since the moment of inertia I of the pendulum is always greater than if the whole mass were concentrated at its center of gravity (see §147); that is, I is greater than Mh^2 and, consequently, l is greater than h.

The point P in line with O and C and at a distance l from O is called *the center of oscillation*. Each portion of the mass of the pendulum between P and O is constrained to swing slower than it would if it were free to oscillate by itself about O as a center, while all portions of the mass below P have to swing more quickly than if they were free. The mass between P and O, therefore, tends to quicken the motion of the pendulum while the mass below P tends to retard it, while the mass situated at P is neither hastened nor retarded, but swings exactly as it would if freely suspended from O.

149. Angular Vibrations. The period of an angular vibration such as that of the balance wheel of a clock is given by

$$T = 2\pi \sqrt{\frac{I}{L_1}} \tag{1}$$

which holds for all cases of a mass vibrating about an axis and acted upon by a restoring moment or couple proportional to the angular displacement of the mass from its rest position. In this formula

T = period of the vibration in seconds.

I = moment of inertia of the body about its axis of vibration (§§ 144–147).

L_1 = restoring couple per radian of angular displacement of the mass from its rest position, in absolute units.

If the moment of inertia of a clock balance wheel is I gram cm^2 and a turning couple of L_1 dyne cm is necessary to twist the balance wheel 1 radian from its rest position, its period of vibration is given by formula (1). This formula is seen to be exactly analogous to that of § 135 and § 811.

150. Center of Percussion. If a rod or pendulum is suspended from an axis A (Fig. 72) and if that axis is given a sudden sidewise impulse or if it is moved rapidly back and forth from side to side, the inertia of the rod will cause it to move as though a certain point B was fixed and the rod turned about that point as axis.

This instantaneous center of the motion is not the center of gravity C, but is the center of oscillation corresponding to the axis of suspension at A. A marble placed on a little shelf at B is scarcely disturbed by the sudden to-and-fro movements of the axis A, while at any other point it would be instantly thrown off.

On the other hand, when the pendulum suspended from the axis A is hanging at rest, if a sudden sidewise impulse is given to the bar at B, as when it is struck a blow at that point, no sidewise impulse is communicated to A in consequence, but the bar simply tends to turn

FIG. 72

about A as an axis. For this reason the point B is also called the *center of percussion* corresponding to the axis A.

In case of a baseball bat the blow is given to the ball with the least jar to the hands when the ball is struck at the center of percussion of the bat corresponding to an axis at the point where it is grasped.

151. Tabulation of Mechanics Formulas. The following is a tabulation of the important formulas studied thus far under mechanics. There is a close parallelism between the formulas for translational and those for rotational motion which is seen by comparing the left hand and right hand columns of the table.

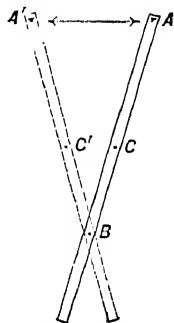| *Translation in Straight Line*	*Rotation about a Fixed Axis*
s = displacement in time t.	θ = angle through which body turns in time t.
v_1 = velocity at *beginning* of interval t.	ω_1 = angular velocity at *beginning* of interval t.
v_2 = velocity at *end* of interval t.	ω_2 = angular velocity at *end* of interval t.
a = acceleration.	α = angular acceleration.

Translation in Straight Line Cont.

$$s = \left(\frac{v_2 + v_1}{2}\right) t \text{ or}$$

$$s = v_1 t + \frac{at^2}{2}$$

$$2as = v_2^2 - v_1^2 .$$

$m = $ inertia or mass.

$F = ma = $ force.

$mv = $ momentum.

$Ft = mv_2 - mv_1 = $ change of momentum.

$\frac{1}{2}mv^2 = $ kinetic energy.

$$T = 2\pi \sqrt{\frac{m}{F_1}} = \text{period of a linear}$$

vibration where $F_1 = $ restoring force per unit displacement from rest position in absolute units.

Rotation about a Fixed Axis Cont.

$$\theta = \left(\frac{\omega_2 + \omega_1}{2}\right) t \text{ or}$$

$$\theta = \omega_1 t + \frac{at^2}{2}$$

$$2a\theta = \omega_2^2 - \omega_1^2 .$$

$I = $ rotational inertia or moment of inertia.

$L = I\alpha = $ torque.

$I\omega = $ angular momentum.

$Lt = I\omega_2 - I\omega_1 = $ change of angular momentum.

$\frac{1}{2}I\omega^2 = $ kinetic energy of rotation.

$$2\pi\sqrt{\frac{I}{L_1}} = \text{period of a torsional vibration about a fixed axis where } L_1 = \text{restoring torque per radian of angular displacement from rest position in absolute units.}$$

PROBLEMS

1. A cylinder weighing 30 kgms. and having a diameter of 1 meter is mounted on an axis and set rotating by a pull of 2 kgms. on a cord wound on an axle 10 cms. in radius. Find the acceleration produced and the speed of rotation 3 sec. from the time of starting. The moment of inertia of a cylinder about its axis is $M\frac{r^2}{2}$ (§ 146) or

$$I = \frac{30 \times 1000 \times 50^2}{2} = 37,500,000 \text{ grm. cm.}^2$$

The force acting is 2 kgms. or 2000 gms. or 2000×980 dynes and the moment of the force is $200 \times 980 \times 10$ dyne-cm.
Substitute in the formula $L = I\alpha$.

$$2000 \times 980 \times 10 = 37,500,000 \alpha \therefore \alpha = 0.523$$

Hence the system will gain in 1 second an angular velocity of a little more than half a radian per sec.
In 3 seconds it will acquire an angular velocity $\omega = 3\alpha = 1.569$, that is, it will be turning at the rate of about 1 revolution in 4 seconds, since $\omega = \frac{2\pi}{T}$ where T is the period of revolution.

2. What is the kinetic energy of a wheel which has a moment of inertia 20 lb. ft.^{2} and is rotating at the rate of two turns per sec.?

3. If a 5-lb. weight is raised by means of a rope wound on the axle of the wheel in problem 2, how high will it be raised before the wheel comes to rest?

4. A uniform rod 40 cms. long and weighing 200 gms. can rotate about a transverse axis through its middle point. How many ergs of work will be required to make it revolve at the rate of three turns per sec.?

5. Suppose the rod in problem 4 is set in rotation by means of a 200-gm. weight attached to a cord wrapped around a cylindrical axle 4 cms. in diameter. How far will the weight have descended in giving a speed of rotation of 3 revolutions per sec.?

NOTE: First solve neglecting the kinetic energy acquired by the 200-gm. weight as it sinks. Then obtain the more exact solution taking account of this energy.

6. The fly wheel of an engine weighs 1200 lbs., the bulk of the weight being in the rim of the wheel at a distance of about 3 ft. from the axis. What is approximately its moment of inertia and how many ft.-lbs. of work must be done by the engine to set it rotating 3 times per sec.?

7. How much energy will be given out by the fly wheel in problem 6 in slowing down from 3 to 2.5 revolutions per sec.?

8. A uniform bar 3 ft. long swings as a pendulum about an axis at one end. Show that the equivalent simple pendulum is 2 ft. long.

9. A uniform spherical steel ball 6 cms. in diameter is hung as a pendulum by a steel wire so that the center of the ball is just 100 cms. below the axis of suspension. Find how far the center of oscillation is below the center of the ball and what is the length of the equivalent simple pendulum, neglecting the mass of the suspending wire.

10. A rectangular bar of steel 1 × 1 × 12 cm. and weighing 90 gms., when suspended in a horizontal position by a wire attached to its middle point, is set oscillating about a vertical axis through its center and makes 4 complete vibrations in 10 sec. Find the moment of force or torque due to the twist in the wire when the bar is at right angles to its equilibrium position.

11. Find the period of oscillation of a solid metal sphere 6 cms. in diameter and weighing 800 gms. when hung by the same wire as the bar in problem 10 and set oscillating about a vertical axis through its center.

Some Cases of Motion with Partly Free Axis

152. Foucault's Pendulum Experiment. It occurred to the French physicist, Foucault, that since a pendulum undisturbed by external forces must persist in its original direction of vibration,

if one were swung at the north pole by some suspension which could not transmit torsion, its direction of vibration would remain constant while the earth turned around under it, so that to an observer moving with the earth the pendulum would seem to change its direction of vibration at the rate of 15° per hour.

At the equator the direction of the meridian remains parallel to itself as the earth rotates, and consequently the plane of vibration of the pendulum would remain unchanged.
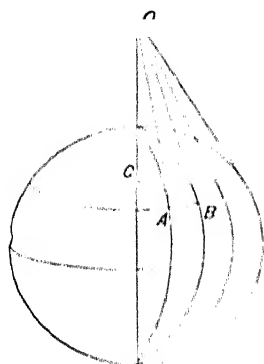

FIG. 73

At any intermediate latitude the tangents to the meridians at two points differing in longitude by 15°, such as A and B (Fig. 73), will meet the axis at O, and the angle AOB measures the change in direction of the meridian per hour. Consequently a Foucault pendulum in that latitude will shift in one hour through an angle equal to AOB. This interesting experiment was carried out by Foucault in 1851. He used as pendulum a massive ball of copper, hung by a wire more than 50 meters long, from the dome of the Pantheon in Paris.

153. Conservation of Angular Momentum. In any body or system of bodies *the total angular momentum of the system cannot be changed by any internal forces:* for suppose A and B (Fig. 74) are two parts of the system which act on each other, since action and reaction are equal and opposite the force on A is equal and opposite to the force on B; and since the distance from the axis to the line of action of the forces is the same for both, the *moments* of the *forces* about the axis will be equal and opposite, so that in the same time they will give equal and opposite angular momenta to the system and consequently the total angular momentum will not be changed.

For example, in the solar system the planets have not only angular momenta about their own axes, but also angular momenta about the common center of gravity of the system. These angular momenta may be represented as vectors and their re-

sultant found from the vector diagram, and neither the direction nor amount of this resultant is changed by any internal forces, such as the attraction of one planet for another or any possible collisions between them.

154. Angular Momentum of Projectiles. A body having angular momentum tends to keep the direction of its axis of revolution constant, and the greater the angular momentum the harder it is to disturb the direction of the rotation; that is, the slower its axis of revolution will change in direction under any given torque.

So the spin of the rifle bullet or shell from a rifled gun causes it to keep pointing in a nearly constant direction as it flies through the air in spite of the tendency of a long bullet to turn sidewise in consequence of air resistance.

155. Motion of a Top. When a rotating body is acted on by forces which tend to turn it about an axis perpendicular to its axis of rotation the effect is to change the *direction* of the axis of rotation without producing any change in the *amount* of the angular momentum about that axis; precisely as when a force acts on a body at right angles to the direction of its linear motion (§ 115) it changes the *direction*, but not the *speed* of the motion.

FIG. 75. Top with point fixed

The motion of a top affords an excellent illustration of this principle. The top in figure 75 is represented as spinning in the direction indicated by the arrow, but in the inclined position shown it is subject to a downward force W due to its own weight acting through its center of gravity G, and the upward pressure of the floor against the point of the top at A. These two forces are equal and constitute a couple which tends to turn the top about an axis DA perpendicular to its axis of revolution. The effect of the couple is to cause a steady change in the direction of the axis of revolution, the upper end of the top moving around in the circle $EFLH$. This change in the direction of the axis of the top may be called its *precessional* motion.

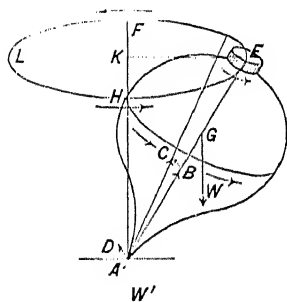The precession of the top may be explained as follows: let the

vector AB represent in amount and direction the angular momentum of the top about its axis, the vector being drawn so that the top is seen to revolve clockwise by an observer looking along the vector AB in the direction in which it points. Similarly the vector AD may represent the angular momentum which would be given to the top in a very small interval of time t by the coup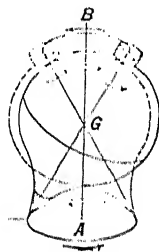le consisting of the forces H and H'. The resultant of the two vectors AD and AB is the vector AC, showing that the resultant angular momentum will have AC as its axis, and the axis of the top will accordingly move through the angle BAC in the time t. And as the vector DA is always at right angles to the plane BAE, the top will move at right angles to this plane, and therefore its upper end E will describe a circle about the vertical axis AK.

FIG. 76

In the case just discussed the friction of the floor is supposed to be sufficient to keep the point of the top fixed at A. But when the top spins on a *frictionless* level surface it remains at a constant inclination and its precessional motion is about a vertical axis through its *center of gravity*, as shown in figure 76.

How it is possible for a top to rise to a vertical position as it spins was first explained by Lord Kelvin. It depends on the fact that the peg of the top is rounded and the friction between it and the floor causes it to roll around in a circle; and when this rolling of the peg on the floor urges the top around faster than the regular precessional motion, it causes the inclination of the top to gradually diminish until it stands vertical, and "goes to sleep." On a perfectly frictionless surface a top could not rise in this way.
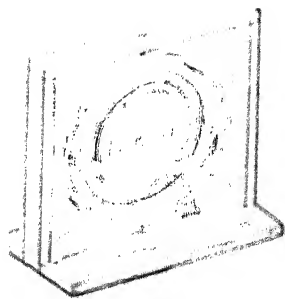

FIG. 77

156. Gyroscope. In the gyroscope shown in figure 77 a wheel with heavy rim is mounted in two pivoted rings so that the axis of rotation of the wheel may be inclined at any angle and the

whole may also turn freely about a vertical axis. When the wheel is in rapid rotation a sharp blow given with the hand to one of the rings as if to change the direction of the axis of rotation, will cause the wheel to vibrate as though it were held in its position by stiff springs.

When a small weight is hung on near one end of the axis of rotation, the wheel, instead of tipping down, rotates slowly around the vertical axis as indicated by the arrow; if the weight is hung from the other end of the axis this precessional motion is reversed. A bicycle wheel serves admirably as a gyroscope.

REFERENCES

A. Gray: "Gyrostats and Gyrostatic Action," Smithsonian Reports, 1914, p. 193.

"The Sperry Stabilizer for Aeroplanes," *Scientific American*, Aug. 8, 1914.

"Gyro-compass," *Scientific American Supplement*, v. 72, p. 200, 1911.

John Perry: *Spinning Tops.*

H. Crabtree: *Spinning Tops and Gyroscopic Motion.*

Worthington: *Dynamics of Rotation.*

157. Precession of the Equinoxes. The earth itself illustrates the precessional motion of the gyroscope. It is a rotating body with enormous angular momentum. But as it is not a sphere and its axis is not perpendicular to the plane of its orbit, the attraction of the sun on the bulging equatorial belt tends to turn it over and make its axis perpendicular to the ecliptic. The effect of this rotational force is a slow precessional motion of the axis of the earth, just as in the gyroscope. The axis remains inclined $23\frac{1}{2}°$ to the pole of the ecliptic, but describes a circle about that pole in a period of about 25,800 years.

If we take the pole of the ecliptic as a center and describe a circle of $23\frac{1}{2}°$

FIG. 78

radius it will pass through the present pole star and will mark the path which is being described by the polar axis of the earth. In about 13,000 years the bright star Vega in the constellation of the Lyre will be very nearly at the pole.
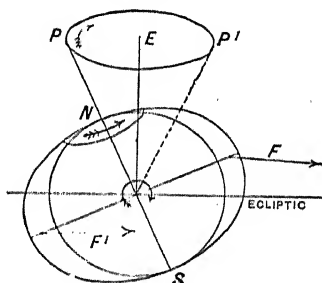
V. UNIVERSAL GRAVITATION

158. Kepler's Laws. The German astronomer, Kepler, in the year 1609, having made a careful study of the observations made

by Tycho Brahé, came to the conclusion that the orbits of the planets were not circular as had been supposed, but elliptical, and announced his discovery in the following laws:

1. *The orbits of the planets are ellipses having the sun at one focus.*

2. *The area swept over per hour by the radius joining sun and planet is the same in all parts of the planet's orbit.* Hence the planet moves faster in its orbit when near the sun than when farther away.

After nine years more of persistent search for some relation between the periodic times of the planets and their distances from the sun, he discovered and announced his third law:

3. *The squares of the periodic times of the planets are proportional to the cubes of their mean distances from the sun.*

159. Newton's Principia. In 1686 Sir Isaac Newton published his great work, the *Principia*, in which he clearly enunciated the fundamental principles of mechanics and applied them to a great variety of important problems. In this work he showed from the laws of mechanics that if the planets moved about the sun in ellipses in the manner described in the first two laws of Kepler, then each planet as it moves in its orbit must be subject to a force which is directed toward the sun, and varies inversely as the square of the distance between them.

160. Universal Gravitation. From the above result Newton concluded that probably all masses, great and small, attract each other with a force proportional to their masses and inversely proportional to the square of the distance between them.

According to this law, the attractive force between any two masses m and M is expressed by the formula

$$F = \frac{mM}{r^2} C$$

where r is the distance between the centers of the masses if they are spherical.* The quantity C is an absolute constant for all kinds of matter and depends only on the units in which force, mass, and distance are measured. It is called the *gravitation*

* According to Einstein's theory of relativity the exact law of gravitation varies slightly from this but the deviation is so small that it does not have to be taken account of in ordinary astronomical calculations.

constant and is equal to the force with which two unit masses attract each other when placed unit distance apart.

161. Moon's Motions Connected with Fall of Apple. Newton conceived that the weight of a body near the surface of the earth is due to this gravitation attraction between the earth and the body, and that an apple drops toward the earth in accordance with the same gravitation law which determines the motion of the moon in its orbit.

To test this point let us, following Newton, find the acceleration which the apple would have if it were dropped toward the earth when as far off as the moon, and compare this acceleration with that which the moon is known to have.

According to the law of gravitation (§ 160), the earth attracts a body at its surface with 3600 times the force that it would if the body were 60 times as far from its center, or at the distance of the moon. Consequently the acceleration toward the earth of a body at the distance of the moon should be $\frac{1}{3600}$ of the acceleration of gravity at the earth's surfaces.

But the acceleration of the moon toward the earth may be computed from the formula

$$a = \frac{v^2}{r} \quad \text{or} \quad a = \frac{4\pi^2 R}{T^2}. \quad (\text{§ 114 and § 118})$$

where R is the radius of its orbit (240,000 miles) in feet and T is its period of orbital revolution (27.322 days) in seconds.

Substituting, we have

$$a = \frac{4\pi^2 \times 240,000 \times 5280}{(2,360,620)^2} = 0.008974 \text{ ft./sec.}^2$$

which is $\frac{1}{3600}$ of 32.30 ft./sec.,^{2}

while the acceleration of gravity at the pole, where it is not affected by the earth's rotation is 32.26 ft./sec.^{2} The two results therefore agree as exactly as could be expected with the data used.

We conclude, then, that the motion of the moon and the fall of an apple or stone are both according to the same law of gravitation.

162. Determination of the Gravitation Constant. To deter mine the constant of gravitation the force of attraction between two known masses must actually be measured. The extreme minuteness of this attraction between small masses makes the exact determination of its value very difficult.

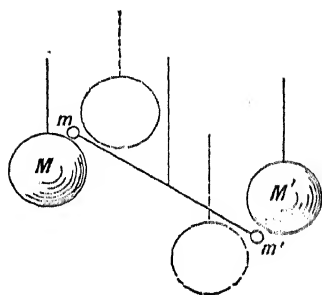It was first accomplished by Cavendish in 1798, using a form of apparatus indicated in figure 79. Two small spherical balls m and m' were mounted on the ends of a light cross bar which was sus pended by a fine silver wire at its center. Two large spherical balls of lead M and M' weighing 158 kilo grams apiece were suspended one near m and the other near m' but on opposite sides so that their attrac tions tended to turn the bar in the same direction. To protect the sus pended bar from being disturbed by air currents it was entirely enclosed in a narrow box, its deflections being observed by a telescope through a glass window.

FIG. 79

Having observed the deflection of the bar when the large masses were in the positions shown, the masses were moved into the dotted positions where their attractions produced a deflection of the bar in the opposite direction. From these ob servations, combined with a measurement of the force required to turn the suspended bar through a given angle, the force of attraction between the masses was determined.

In the year 1889 C. V. Boys, who had discovered the remark able elastic properties of fine quartz fibers, devised an apparatus similar in principle to that of Cavendish, but much more com pact, in which the small suspended masses were hung by a quartz fiber so fine that larger deflections and greater accuracy of meas urement were attained.

According to Boys' determination, $C = 6.6576 \times 10^{-8}$ in C. G. S. units. That is, the attraction between two masses of one gram each concentrated at two points a centimeter apart, or of two spherical masses of one gram each with a distance of one centimeter between centers, is 0.000,000,066.6 dyne.

Two kilogram masses 10 cms. between centers attract with a force of 0.000666 dyne, or about seven ten-millionths of a gram weight.

The constant of gravitation has also been reckoned by estimating the mass contained in an isolated mountain and then measuring its deflecting effect on a plumb-line near its base.

163. Mass of the Earth. When the gravitation constant is known the mass of the earth itself may readily be determined. For consider the earth as attracting a gram mass at its surface. The force of attraction is g dynes or approximately 980, and from the law of gravitation

$$F = \frac{mM}{r^2} C.$$

Take M = mass of the earth, F = 980, m = 1, r = radius of earth in centimeters, and C = 6.66×10^{-8}.

All of these quantities are known except M, which may be calculated. In this way the mean density of the earth is found to be 5.527, a result which is especially interesting as the average density of the *surface* materials of the earth is only about 2.5.

164. Mass of a Planet. So also the mass may be found of any planet having a satellite whose distance and period of orbital revolution about the planet can be observed. For the attraction between the planet and satellite is expressed by $\frac{mM}{r^2} C$, while the centripetal force in case of a satellite of mass m and period T and moving in a circle of radius r, is $\frac{4\pi^2 m}{T^2} r$, and since it is the attraction which holds the satellite in its orbit we have

$$\frac{mM}{r^2} C = \frac{4\pi^2 m}{T^2} r.$$

In the equation the mass of the satellite m cancels, and as all the other quantities except M are known, the mass of the planet may be computed.

165. Significance of Kepler's Third Law. Let M represent the mass of the sun, E the mass of the earth, r the mean distance between them, and T the period of the earth's revolution about the sun. Then, as in the last paragraph

$$\frac{ME}{r^2} C = \frac{4\pi^2 Er}{T^2} \quad \text{or} \quad \frac{MC}{4\pi^2} = \frac{r^3}{T^2}. \tag{1}$$

So also if J is the mass of some other planet, such as Jupiter, and if r_1 and T_1 represent its distance from the sun and period of revolution in its orbit, respectively, we have

$$\frac{MJC}{r_1^2} = \frac{4\pi^2 Jr_1}{T_1^2} \quad \text{or} \quad \frac{MC}{4\pi^2} = \frac{r_1^3}{T_1^2}$$

If the constant of gravitation C has the same value in case of the sun and earth as it has in case of the sun and Jupiter, then

$$\frac{r^3}{T^2} = \frac{r_1^3}{T_1^2}$$

which is precisely what Kepler's third law asserts to be true throughout the solar system. It is concluded, therefore, that the same gravitation constant holds everywhere throughout the solar system and probably throughout the material universe.

166. Variation of Gravity on Earth. The force of gravity is not the same everywhere on the earth's surface. There are three circumstances which determine this variation, namely, the fact

FIG. 80

that the earth is not a sphere, its rotation, and the height above sea level of the given station.

The earth is approximately an oblate spheroid having its polar radius less than its equatorial by 13.2 miles or 21.2 kilometers and in consequence of this the value of g at the poles is greater than at the equator by 1.6 cm./sec.², due to this cause alone. But there is another circumstance which still further reduces the value of g at the equator. The *rotation of the earth* affects both the direction and amount of the acceleration g. For the resultant attraction F of the earth on a gram of matter situated at A (Fig. 80) is directed toward the center O, but this resultant attraction serves both to supply the centripetal force f, which holds the mass on the earth as it rotates, and also the force which we call its weight which gives it acceleration g when dropped. The centripetal acceleration f is directed perpendicular to the polar axis and is equal to $\frac{4\pi^2}{T^2}r$, where T is the period of rotation of the earth and r is the distance AB.

The distance $AB = R \cos l$ where R is the radius of the earth and l the latitude of A. Evidently then, f is a maximum at the equator and has zero value at the poles. Since F is the resultant of f and g, and is directed toward the center of the earth, it is clear from the diagram that g cannot be directed toward the earth's center except at the poles or equator. The direction of g is the direction in which a plumb-line will hang or a body will fall at A. Also a liquid surface, as the surface of the ocean, must be at right angles to g (see § 179).

At latitude 45° the plumb-line points away from the center of the earth about 6.9 miles.

At the equator the centrifugal force of a mass of one gram is 3.36 dynes. Hence the acceleration of gravity is less at the equator than at the poles by 3.36 cm./sec.² on this score alone.

The height of a place above sea level also affects the value of g, as it must diminish with the increase in distance from the center of the earth. If h represents the height in centimeters or in feet, the corresponding change in g is $(0.000003)h$.

Though on account of the irregular shape and distribution of the earth's mass the exact value of g at any place can be determined only by pendulum experiments, an approximate value may be calculated for any place on earth by the following formula due to Clairaut:

$$g = 980.6056 - 2.5028 \cos 2\lambda - 0.000003h.$$

where λ represents the latitude of the place and h its height above sea level.

SOME VALUES OF G AT SEA LEVEL

PLACE	CM./SEC.²	FT./SEC.²	PLACE	CM./SEC.²	FT./SEC.²
Pole...........	983.1	32.25	New York.....	980.2	32.16
London........	981.2	32.19	Washington....	980.0	32.15
Paris..........	980.9	32.18	Equator.......	978.1	32.09

MECHANICS OF LIQUIDS AND GASES

PART I. FLUIDS AT REST

167. Fluids. Certain substances, such as air, water, glycerin, etc., are characterized by great mobility, changing their shapes and flowing under the smallest forces. They are known *fluids*

Fluids are divided into two classes, liquids and gases.

Liquids change but slightly in volume when sub ject to great pressure and may have a free surface.

Gases are far more compressible than liquids and fill parts of the containing vessel. Water is a type of liquid, and air of gas.

168. Density. *The mass of any substance contained in unit volume is known as its density.* In the C. G. S. system of units density is expressed in grams per cubic centimeter, while in the foot-pound-second system it is expressed in pounds per cubic foot.

Thus the density of water is 1.0 on the first system, while it is 62.5 on the latter system.

A table showing the densities of some substances will be found on page 158.

169. Viscosity. Fluids differ greatly in mobility. If a dish of water is tilted, the flow is so rapid that it gives rise to waves that surge to and fro, while in case of glycerin or syrup the flow is slow and the liquid only gradually settles to the new level. This difference in mobility is due to *viscosity* or internal friction (§ 254). Substances like pitch or tar are very viscous, while water, alcohol, and ether are but slightly so.

A *perfect* fluid is one that has no viscosity and is an ideal. All known fluids, even gases, have some viscosity.

170. Force in Fluid at Rest. *The force exerted by a fluid at rest against any surface is perpendicular to that surface.* Otherwise, owing to the mobility of the fluid, flow must take place along the surface, which of course cannot be in a liquid at rest.

This law is true of all fluids, even those which are very viscous, *after they have settled into equilibrium.*

171. Pressure. Let a very small flat surface be imagined at some point in a fluid. The fluid on one side of that surface exerts a force perpendicular to the surface against the fluid on the opposite side. This force is proportional to the surface, and *the force per unit surface is called the pressure.*

In C. G. S. units pressure is measured in dynes per square centimeter; it may also be measured in grams per square centimeter, pounds per square inch, etc.

172. Hydrostatic Pressure. *At any point in a fluid at rest the pressure is the same in every direction.* This is a direct consequence of the mobility of fluids, for a little sphere of liquid at the given point could not be in equilibrium if the pressure against its surface were not the same in every direction.

173. Pressures on Same Level. *In a liquid at rest the pressure is the same at all points on the same level.* For a horizontal cylindrical column of liquid reaching from A to B is in equilibrium under the pressure of the surrounding liquid. The pressure against its sides is perpendicular to the line AB, and therefore has no influence to move the column toward A or B. And since it is level it has

FIG. 81

no tendency to slide toward A or B by reason of its weight. The force against the end at A must therefore be balanced by the force against the end at B. These forces are due to the pressures at A and B, and since the ends have equal areas the pressure at A must be equal to the pressure at B.
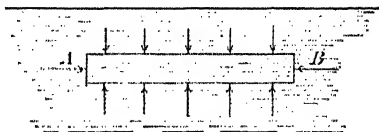
174. Pressures at Different Depths. *The difference in pressure between two points at different levels in a mass of fluid at rest under gravity is equal to the weight of a column of the fluid of unit cross section reaching vertically from one level to the other.* For a vertical cylindrical column of the fluid of unit cross section reaching from B to C is in equilibrium under the pressure of the surrounding fluid. The pressure against the sides of the vertical column is horizontal and has no power to support its weight, consequently the upward force at C must balance the weight of the column in addition to the downward force at B. Hence, since the force against the end of a unit column is equal to the

pressure, the pressure at C is greater than the pressure at B by the weight of the column of fluid of unit cross-section reaching from B to C.

If h is the height of the column in centimeters and d is the weight of one cubic centimeter of the fluid in grams, then hd is the weight of the column and is thus the difference in pressure between B and C in grams per sq. cm.
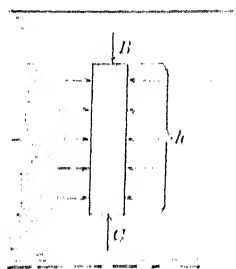
FIG. 82

The difference in pressure expressed in dynes per sq. cm. is hdg where g is the acceleration of gravity in cm. sec.² The total pressure at a point h centimeters below the surface is therefore as follows:

Pressure in grams per sq. cm. = hd + pressure on surface in grams per sq. cm.

Pressure in dynes per sq. cm. = hdg + pressure on surface in dynes per sq. cm.

Note as to Units. In calculating pressure by the use of the formula hd, it must be remembered that if the pressure is to be found in pounds per square inch, then h must be expressed in inches and d is the weight of one cubic inch of the liquid in pounds. The student is advised, however, to compute directly the weight of a column of the substance of unit cross-section without thinking of any formula.

In gases the density is so small that the pressure is practically the same everywhere throughout a *small* volume.

Pascal's Principle. *Pressure is transmitted equally in all directions throughout a mass of fluid at rest, or if the pressure at any point is increased, it is increased everywhere throughout the fluid mass by the same amount.*

175. Hydraulic or Hydrostatic Press. An important mechanical device known as the *hydraulic press* is a good illustration of the application of the laws of fluid pressure. It was first constructed by Bramah in 1796, and is sometimes known as Bramah's press.

It consists of a strong cylinder in which works a cylindrical piston or ram of larger diameter. A collar of oiled leather or copper surrounds the piston in such a way that the greater the pressure of the liquid filling the cylinder, the more closely does the collar fit the piston. By means of a small pump, oil or water is forced into the large cylinder, a check valve preventing its return. In consequence of the law of pressure just enunciated,

whatever pressure is communicated to the liquid by the pump
will be exerted everywhere equally against the walls of the con-
taining cylinders. So that if the large piston has 100 times the
area of the other it will exert a force 100 times as great as that
applied to the pump piston.

Hydraulic jacks act on this principle: they contain a reser-
voir of oil which may be pumped into the main cylinder, thus
forcing up the ram; opening a
small stopcock permits the flow
of oil back to the reservoir. Oil
is used as it keeps the machine
lubricated and does not freeze.

It is to be observed that when
the liquid in the hydraulic press
is incompressible as much work
is done by the large piston as is
expended upon the smaller one.

**176. Pressure Independent
of Shape of Vessel.** It has been
shown that the pressure at any
point in a liquid under gravity
depends only on the depth of the
point below the surface, on the

FIG. 83.   Hydrostatic press

density of the liquid, and on the pressure on its surface.

The total force exerted against the bottom of a vessel by the
pressure of the liquid which it contains is the product of the
pressure at the bottom by its area, and may therefore be very
different from the actual weight of liquid which the vessel con-
tains; and when a vessel is filled with water to a given height the
force against its bottom is the same whether the upper part of
the vessel is flaring, cylindrical, or narrow. The reasonableness
of this result will be evident from the following considerations.

In the case of the vessel with flaring sides we may think of a
cylindrical column resting on the bottom and pressed upon by
the surrounding water as shown in the figure (Fig. 84). This
pressure is necessarily perpendicular to the surface of the cylin-
drical column and, therefore, can have no effect in either support-
ing it or pressing it down. The whole weight of the cylindrical
column is, therefore, supported by the bottom plate. In case of
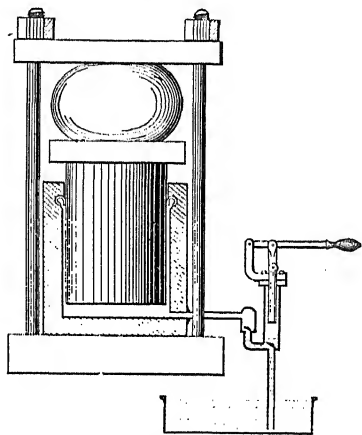
the vessel which is narrow at the top, the liquid exerts a downward force on the bottom greater than its weight because the sides of the vessel press the liquid down, just as a man in a box
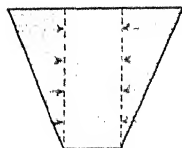
FIG. 84

may brace himself against the top and press against the bottom with a force far greater than his own weight.

This fact that the force exerted on the bottom of a vessel may be greater than the weight of all the liquid in the vessel has been called the *hydrostatic paradox*.

Pascal succeeded in bursting a strong cask by the pressure produced by a column of water in a narrow pipe 40 ft. high.

LIQUID SURFACES

177. Free Surface of a Liquid. When a liquid is at rest or in equilibrium *the force which a surface particle exerts against the adjoining liquid must be perpendicular to the free surface at that point*, otherwise the particle would move along the surface. This force depends upon gravity, on the attraction of neighboring particles, and on the atmospheric pressure on the surface, and also upon any acceleration which the particle may have.

178. Level Surface. When a liquid is at rest on the earth, all parts of the surface which are not too near the walls of the containing vessel are at right angles to the direction of gravity or to the direction in which a plumb-line points. Such a surface is called *level*. A level surface is not a *flat* surface, but has the same curvature as the earth. In a pond 1 mile in diameter the center is 2 in. higher than a plane passing through the edges.

The force is not necessarily the same at all points of a level surface. This is well illustrated in case of the earth, for the

force of gravity at sea level near the poles is decidedly greater than at the equator.

179. Surface of a Rotating Liquid. When a vessel containing a liquid is rotated by a whirling machine, the liquid by virtue of its viscosity soon comes into equilibrium,* and turns at the same rate as the vessel. If the speed is slow the upper surface of the liquid is slightly concave, at greater speed it will become deeply hollowed, but it always has the form of a paraboloid of revolution. Here a little mass m exerts against the adjoining liquid a downward force mg due to gravity, and an outward centrifugal force † equal to $m\omega^2 r$. The components q due to gravity (Fig. 86) are the same at all points of the surface,

FIG. 86.   Surface of rotating liquid

while the centrifugal components l_1, l_2, l_3 increase in proportion to the distance of the particle from the axis of rotation. The resultant forces a_1, a_2, a_3 will therefore be differently inclined, and
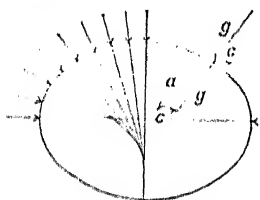
FIG. 87

the surface must be of such a curve as to be at right angles to them. It will be noted that the resultant force is greater at points higher up on the surface, so that a surface particle near the top presses against the surrounding liquid with far more force than it would if at the bottom of the curve.

The oblate form of the earth is similarly explained. A unit mass at the earth's surface exerts a downward force a toward the center of the earth due to attraction, and also a centrifugal force c due to rotation (Fig. 87). The latter component is zero at the poles and reaches a maximum at the equator and is always at right angles to the polar axis. The resultant downward force g is, therefore, directed exactly toward

* That is, it is in equilibrium considered as a whole, though the individual particles move in circles and are therefore accelerated.

† The pressure of the adjoining parts against any little liquid mass supplies the *centripetal* force urging it toward the axis as it rotates. Its outward reaction against that pressure is the *centrifugal* force.

the center only at the poles and at the equator, and the surface of the ocean when calm must be everywhere perpendicular to g.

180. Surface in Connected Vessels. In a continuous mass of one kind of liquid all points on the same level must   at the same pressure, even though they may be in separate branches

FIG. 88

of the containing vessel. Thus the pressure at B (Fig. 88) is the same as at B', and that at C is the same as at C''. It is clear that the enclosed air is under greater pressure than that of the atmosphere at A.
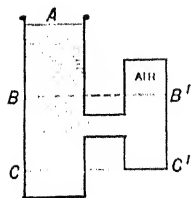
When communicating parts of a vessel of liquid are open to the air the free surfaces must lie all on the same level because all are at the same pressure.

181. Case of Two Liquids. If a bent tube containing mercury, as shown in the figure, have some other liquid, as water or oil, poured into the longer arm, the mercury will be pressed down on that side and raised on the other. Since all below A is one continuous liquid, the pressure at A must be the same as at A' on the same level, hence the column of mercury BA' must produce the same pressure as the column of liquid CA.

Letting h and h' represent the heights of the two columns of liquid and d and d' their densities, then, since the pressures of the two columns must be equal,

$$hd = h'd'.$$

182. Spirit-level. The ordinary spirit level consists of a glass tube hermetically sealed, nearly filled with alcohol or ether, a

FIG. 89

FIG. 90.   Spirit-level

bubble of air or vapor being left. The tube is bent slightly, forming the arc of a large circle, and the bubble always rests in equilibrium at the highest point.

A level is said to be *sensitive* when a small inclination will cause a large motion of the bubble. In a sensitive level the curvature of the tube is very slight, and the bubble is usually
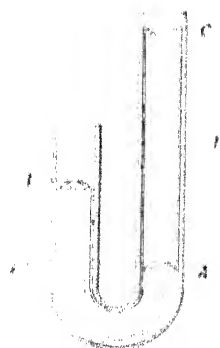
large, otherwise it would be sluggish in its movements. For fine levels the tube is carefully ground on the inside so as to have a uniform curvature.

PROBLEMS

1. Find the pressure 3.50 meters below the surface in a pond of water; in grams per sq. cm. and in dynes per sq. cm.

2. Find the pressure in pounds per sq. in. 30 ft. below the surface of a pond, taking the weight of 1 cu. ft. of water as 62.5 lbs.

3. A piston 1 ft. in diameter carries a weight which together with that of the piston amounts to 200 lbs. How high a column of water will be required to produce enough pressure under the piston to support the weight?

4. What is the pressure 1 mile below the surface of the ocean, in pounds per sq. in., taking the relative density of sea water as 1.03?

5. Find the difference between the pressure at the bottom of a vessel 75 cms. deep filled with water, and the pressure when the vessel is full of mercury. Density of mercury = 13.6.

6. A jar has a square cross section 5 cms. each way and is 30 cms. deep. It is half full of mercury and half full of water; find the pressure halfway down and also at the bottom, also the total force due to pressure against the bottom.

7. Find the total force against one side due to pressure in the preceding problem.

8. If a cubical tank 4 ft. each way is level full of water, find the pressure in pounds per sq. in. on bottom. Also the total force against one side in lbs. weight.

9. Oil of density 0.7 is poured into one branch of a U-tube which contains enough mercury to keep the bend full. When the column of oil is 39 cm. high, how much higher will it stand than the mercury in the other branch?

10. When the atmospheric pressure is just 1,000,000 dynes per sq. cm., how far below the surface of a pond of water will the total pressure be just twice as much as at the surface?

11. In a pail of water spinning about a vertical axis through its center the surface of the water is hollowed so that at a point 10 cms. from the axis the surface is inclined 45°. Find the number of revolutions per sec. which the pail is making.

BUOYANCY AND FLOATING BODIES

183. Buoyant Force of a Fluid. Suppose that a mass of wood or iron is immersed in a liquid and it is required to find the force exerted upon it by the surrounding liquid. Imagine the given

substance removed and its place filled by the liquid, and conceive of this portion as separated from the surrounding liquid by an imaginary surface *ABC* of the same shape as the original body. The liquid is in equilibrium, and since the mass enclosed
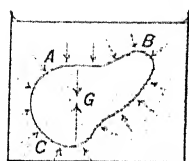
FIG. 91

in the surface *ABC* is urged down by its own weight, this weight must be exactly balanced by the force due to the pressure of the surrounding liquid on the surface *ABC*. Hence the resultant force due to pressure on the surface is an upward force equal and opposite to the weight of the enclosed mass of liquid, and since the whole weight of the enclosed mass acts down through its center of gravity *G*, the *center of pressure* must also be at the same point.

Now, neither the amount nor direction of the pressure will be changed at any point of the surface *ABC* if it is filled with wood or iron instead of the liquid. Therefore *when any object is wholly or partially immersed in a liquid it is buoyed up by a force equal to the weight of the displaced liquid, and the center of pressure is where the center of gravity of the submerged portion would be if it were homogeneous.*

There is nothing in the above reasoning which restricts this conclusion to *liquids*, it may therefore be stated as a general law of *fluids* and is known as *Archimedes' principle*, from its discoverer.

184. Experimental Illustration. A brass cylinder which exactly fits into and fills a cup is suspended together with the cup from one pan of a balance and exactly counterpoised by weights. A vessel of water is raised under the cylinder until it is quite immersed, and the weights will now greatly overbalance the

FIG. 92

cup and cylinder; but if the cup is just filled with water the balance is restored.
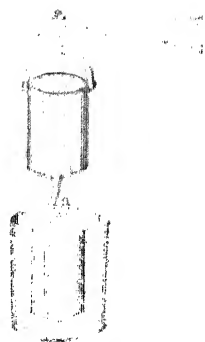
185. Buoyancy at Great Depths. Since buoyant force depends on the weight of the liquid displaced and not directly on the pressure, it makes no difference whether the immersed body is 1 in. or 100 ft. below the surface of the liquid except for the

compression due to increased pressure. If the immersed body is more compressible than the surrounding liquid it will displace less liquid where the pressure is great than at the surface and so will be less buoyed up at great depths. If it is less compressible than the liquid, it will be more buoyed up at great depths than when near the surface.

The heavy iron shot used in deep sea soundings is buoyed up slightly more at great depths than at the surface because water is more compressible than iron.

186. Cartesian Diver. The Cartesian diver is a small bulb of glass open at the bottom and containing just enough air to cause it to float in a jar full of water. A sheet of rubber is tied firmly over the mouth of the jar, and by pressing on the rubber the pressure in the liquid is increased and the air in the bulb compressed into smaller volume. The bulb with the contained air may thus be made to displace less than its own weight of water and will then sink to the bottom, but rises again when the pressure is relieved and the air expands.

187. Equilibrium of Floating Bodies. A floating body may be considered as acted on by two forces; its own weight acting down through its center of gravity and a buoyant force equal to the weight of the displaced liquid acting up through the center of pressure. It can be in equilibrium only when these two forces are equal and opposite. The conditions for equilibrium may then be thus stated:

FIG. 93
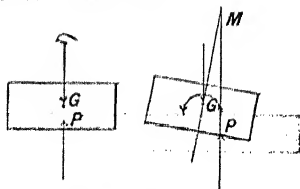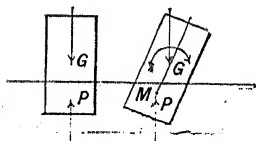
1. *The weight of the displaced liquid must be equal to the weight of the floating body.*

2. *The center of gravity of the floating body must be in the same vertical line as the center of pressure.*

FIG. 94   FIG. 95

The displacement of a ship is the weight of water which it displaces, and is therefore the total weight of the ship and equipment.

188. Stability of Equilibrium. If, when a floating body is slightly inclined from its position of equilibrium, the couple resulting from its own weight and the buoyant force of the liquid tends to turn it back into its original position the equilibrium is said to be *stable*. In figure 96 G is the center of gravity and P the center of pressure of the floating block. When it is tipped slightly P is displaced to one side in such a way that the combined action of the forces through G and P tends to turn the body in the direction of the arrow, bringing it back into its original state of equilibrium which is therefore *stable*. In figure 95 is shown a state of equilibrium such that when the body is slightly displaced the couple acts to increase the displacement and to turn the body away from its original position. In this case the equilibrium is *unstable*.
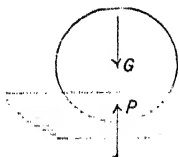
FIG. 96

A floating homogeneous sphere may be turned in any way and the center of pressure P will always be directly under the center of gravity, and the equilibrium will remain undisturbed. Here the equilibrium is

SPECIFIC GRAVITY AND ITS MEASUREMENT

189. Specific Gravity. The relative density of a substance as compared with some standard substance is known as its *specific gravity*. Solids and liquids are usually compared with water as a standard, while gases are often referred to air or hydrogen.

The specific gravity of a substance referred to water is found by dividing the weight of the given substance by the weight of an equal volume of pure water at the temperature of $4°$ C.

The specific gravity of a substance is a *ratio* and is therefore the same whatever system of units is employed.

Since 1 c.c of pure water at $4°$ C.

FIG. 97

has a mass of 1 gram, the density of a substance in *grams* per cubic centimeter is equal to its specific gravity referred to water.

190. Specific Gravities by Balance. The substance, of which the specific gravity is to be determined, is suspended by a fine
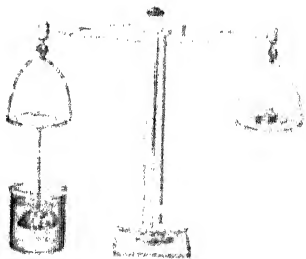
fiber from one arm of a balance and weighed, first in air and then when immersed in water. The second weighing will be less than the first by the weight of the water displaced by the substance. The difference between the two weighings will then give the weight of a mass of water of the same volume as the substance, and therefore if the weight in air is divided by the difference between the weights in air and water the *specific gravity* is obtained.

191. Mohr's Balance. A convenient balance for determining the specific gravity of liquids is that shown in figure 98. A glass bulb weighted so as to sink in liquids is hung from one arm of a balance and exactly counter-
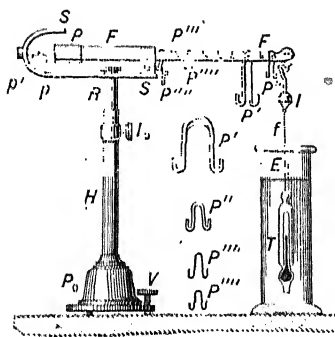
FIG. 98   FIG. 99

poised by the weight *P* on the other arm. The glass bulb is hung in the liquid to be examined and the buoyant force of the liquid balanced by riders hung on the balance arm. From the weight and position of the riders the specific gravity of the liquid is obtained directly without calculation; for the several riders are so adjusted that each has one-tenth the weight of the next larger, and the position of each on the balance arm gives the figure for the corresponding decimal place in the result.

192. Hydrometers of Constant Weight. These instruments are usually made of glass and consist of a rather long light bulb having a slender stem above and a weighted bulb below so that the instrument floats in a vertical position in the liquid whose density is to be determined. By means of a scale on the stem

the specific gravity of the liquid may be read directly from the point on the scale to which the instrument sinks.

In such a case the weight of the whole hydrometer must be equal to the weight of the displaced liquid, so that it v is the volume of the hydrometer below the mark to which it sinks in a given liquid and if d is the weight of unit volume of the liquid, then $W = vd$ where W is the weight of the hydrometer.

The specific gravity scale of a hydrometer is not a scale of equal parts, corresponding divisions being farther apart at the upper end of the stem than at the lower. The Beaumé scale is an arbitrary scale of equal parts in which hydrometers are often graduated.

Hydrometers are made for liquids lighter than water and also for liquids heavier than water.

If the stem of a hydrometer is slender (compared with the volume of the immersed portion of the instrument), it will be sensitive and a small change in density will cause a large change in its immersion.
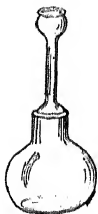
193. Specific-gravity Bottle. When the specific gravity of a powdered substance is to be determined, the specific-gravity bottle may be used. This is a small flask having a carefully ground tubular stopper. The powder to be experimented upon, after being weighed, is put into the flask which is then filled with water up to a certain mark on the stopper. The weight of the whole is then determined and also the weight of the flask when filled with pure water alone up to the same mark. From these three weighings the weight of water displaced by the powder may be determined, and so its specific gravity may be obtained. If the powder is soluble in water some liquid in which it is insoluble must be used.

FIG. 100

PROBLEMS

1. A piece of metal weighs 300 gms. in air and 260 gms. in water. What is its volume, specific gravity, and density?

2. A certain block of wood has a volume of 80 c.c. and specific gravity 0.8. Find the volume and weight of water which it will displace when floating.

3. A certain body weighs 240 gms. in air, 160 gms. in water, and 140 gms. in another liquid. Find the specific gravity of the body and also of the second liquid.

4. A block of wood floats in water with $\frac{3}{7}$ of its volume above the surface. What is its density?

5. A block of wood floats in water with $\frac{1}{5}$ of its volume above the surface, but when floating in oil $\frac{5}{6}$ of its volume is submerged. Find the specific gravity of the wood and of the oil.

6. A piece of metal weighs 16 gms. in air and 14 gms. in water. Another substance B weighs 8 gms. in air, and the two when fastened together weigh 2 gms. in water. Find the specific gravity of each.

7. A sinker weighing 38 gms. is fastened to a cork weighing 10 gms. and the two together are in equilibrium when immersed in water. Find the specific gravity of the sinker if that of the cork is 0.25.

8. The stem of a hydrometer is graduated upward from 0 to 100 in equal parts, and the volume of the instrument below the zero of the scale is three times that of the graduated stem. When placed in water it sinks to the 20-mark. Find the density of a liquid in which it sinks to the 80-mark, also of a liquid in which it sinks to the 0-point.

9. A mass of 80 gms., having a density 8, balances a mass of 140 gms. when both are suspended in water from the arms of a balance. Find the density of the larger mass.

GASES AND ATMOSPHERIC PRESSURE AND BUOYANCY

194. Gases. The second great division of fluids is that of compressible fluids or gases. Those mechanical properties of liquids which are due simply to their fluidity are also possessed by gases. The compressibility of gases, however, is so great that their changes of density due to variations in pressure cannot be neglected.

195. Density of Gases. Gases possess weight as is shown by the following experiment. Pump the air out of a globe of glass or of metal until it is well exhausted, suspend it from one pan of a balance and weigh it. Now open the stopcock in the globe, admitting air, and when it is full weigh it again. The difference between the two weights is the weight of the globe full of air. At temperature 0° C. and when the barometric pressure is 76 cms., a cubic foot of dry air weighs about $\frac{1}{13}$ lb. or a cubic liter 1.293 gms.

A room 30 ft. long, 30 ft. wide, and 10 ft. high contains under ordinary conditions about 700 lbs. of air.

The following table gives the densities of some familiar gases at 0° C. and a pressure of 76.0 cms. of mercury.

GAS	DENSITY IN GMS. PER C. C.	RELATIVE DENSITIES
Air.............	0.001293	1.0000
Oxygen.........	0.001430	1.1057
Nitrogen........	0.001257	0.9720
Hydrogen.......	0.00008988	0.0695
Chlorine........	0.003133	2.43
Carbon dioxide...	0.001974	1.527
Carbon monoxide.	0.0012504	0.967
Ammonia........	0.000761	0.589

196. Torricelli's Experiment. The first measurement of the pressure of the atmosphere was made in 1643 by Torricelli, a pupil of Galileo.
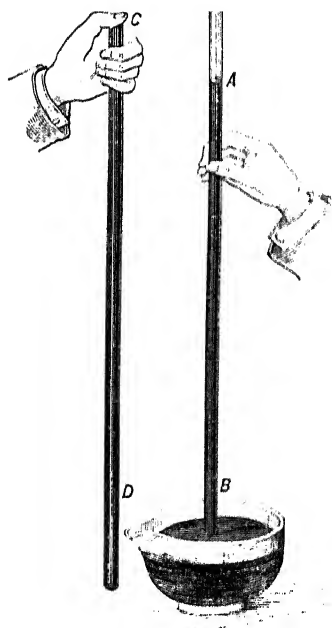
FIG. 101

It was known that water could not be raised more than 34 ft. by a suction pump. Torricelli believed that this was because water was raised in such a pump by the pressure of the atmosphere. He concluded that as mercury was 13.6 times as dense as water the atmospheric pressure would be able to support a column of mercury only ¹⁄₁₃.₆ times as high, or about 30 in. in length, and to test it tried the following experiment.

A tube nearly 3 ft. long and closed at one end was filled with mercury and then the open end being closed with the finger to prevent the escape of mercury the tube was inverted and placed with its open end below the surface of mercury in a dish, after which the finger was withdrawn. The mercury at once sank in the tube till it stood at a height of about 30 in. or 76 cms. above

the level in the dish. The space above the mercury in the tube was a vacuum except for the presence of mercury vapor.

As 1 c.c. of mercury weighs 13.6 grams, the atmospheric pressure able to support a column 76 cms. high must be $76 \times 13.6 = 1033.6$ gms. per sq. cm., and would, therefore, sustain a column of water 1033.6 cms. high, or 33.9 ft.

Pascal, reasoning that if the pressure of the atmosphere was due to its weight the pressure should be less on top of a mountain than at its base, caused the experiment to be tried and established the fact.

197. Magdeburg Hemispheres. Otto von Guericke, of Magdeburg, shortly after he had invented the air pump, demonstrated the pressure of the atmosphere by means of two hemispherical cups of copper carefully fitted together to form a spherical vessel about 2 ft. in diameter. When the air was exhausted from the vessel two teams of horses were unable to pull the cups apart. The force with which the cups are pressed together in such a case is found by multiplying the area of the circular opening of the cups by the difference between the air pressure on the inside and outside.

198. Barometer. Instruments for the measurement of the atmospheric pressure are known as barometers. The best barometers usually employ a column of mercury, as in Torricelli's experiment.

A form much used is the Fortin barometer, the reservoir of which is shown in the figure. The tube containing the mercury is sheathed with brass to protect it from injury, the height of the column being read through an opening

Fig. 102. Fortin barometer

by means of a vernier which slides on a scale graduated on the brass sheath. As the mercury sinks in the barometer tube it flows out into the vessel at the bottom and raises the level there, it is therefore necessary to provide some means of adjusting the height of the mercury in the lower vessel. This is accomplished by the screw C, on turning which the flexible leather bottom of the vessel is raised or lowered until the surface of the mercury

exactly touches the ivory point P, which is the zero point from which the scale is graduated. As the lower vessel is not air tight, the external air pressure is freely transmitted to the surface of the mercury. The greatest care is taken in filling such a barometer that no air is left clinging to its sides, the mercury being usually heated and even boiled in the tube.

199. Capillary Correction. The upper surface of the mercury column in a barometer tube is rounded upward in a *meniscus*, higher at the center than at the edges, and the height of the barometer is measured to the highest point of this curved meniscus.

The effect of the curvature is to make the column stand slightly *less* than if the surface were flat. Hence to obtain the true height a small correction, called the *capillary correction*, which depends on the curvature of the surface, must be added to the apparent height.

In a standard barometer the tube should be so large (2 cms. in diameter) that there is no curvature at the center of the surface in which case there is no capillary correction.

Capillary Correction Millimeters

| Capillary Depression | 1.4 | 0.8 | 0.5 | 0.3 | 0.2 mm |
| Internal Diameter of Tube | 4.0 | 6.0 | 8.0 | 10.0 | 12.0 mm |

200. Temperature Correction. It must be remembered that the scale by which the height of a barometer is read is correct at only one temperature, and also that the density of the mercury itself varies with the temperature; in order, therefore, that barometer readings may be definite, what is known as the *reduced reading* is always given; this is the height at which it would stand if the mercury had the density which it has at 0° C.

Effect of Gravity. It might be supposed that if the reduced heights of the barometers at two places were the same the atmospheric pressures at those places would be equal, but this is not necessarily so. The pressure in grams per square centimeter would be the same, but the weight of a gram depends on the force of gravity. Near the equator a gram weighs 978 dynes, while near the poles it weighs over 983 dynes. If the reduced height of the barometer in centimeters be multiplied by the density of mercury at 0° C. and the product by the acceleration of gravity at the given place, the pressure recorded by the barometer will then be determined in dynes per square centimeter, which is absolutely definite.

201. Aneroid Barometer. An exceedingly convenient and portable form of barometer is known as the *aneroid* (from the Greek, meaning *without liquid*). A disc-shaped metal box, like a small blacking box, is provided with a top made of thin metal corrugated so as to be extremely flexible. The air is exhausted from the box and it is permanently sealed, the top being supported by a stout steel spring which prevents it from collapsing. As the atmospheric pressure increases the spring yields a little and its point moves downward, acting by means of levers and a delicate chain to give a greatly increased motion to the pointer which moves over a graduated dial. A hair-spring serves to take up the

FIG. 103. Diagram of mechanism of aneroid barometer

slack of the chain. Such an instrument may be made as compact and portable as a watch. It is subject to change, however, and needs to be compared with a mercurial barometer from time to time. Also the elasticity of the spring varies with the temperature.

202. Standard Atmospheric Pressure. It is customary in stating the densities of gases to give them at what is called atmospheric pressure. This *standard atmospheric pressure*, sometimes called a *pressure of one atmosphere*, is the pressure of a column of mercury 76 cms. high at 0° C.

When the acceleration of gravity has the value that it has at Paris (980.94) this pressure is 1,013,600 dynes per square centimeter.

At London its value is 1,013,800 dynes per square centimeter.

A unit of pressure frequently used in recording very low pressures is the *bar*. A bar is defined as one-millionth of the pressure corresponding to 75 cm. of mercury at 0° C., being almost equal to one-millionth of the standard atmospheric pressure.

203. Buoyancy. The law of buoyancy, known as Archimedes' principle, that bodies immersed in a fluid are buoyed up with a force equal to the weight of the displaced fluid, holds for gases

as well as for liquids. This may be easily illustrated by the apparatus shown in the figure. A hollow globe is balanced by a solid mass of lead or brass hung from the other arm of the balance. When the globe is closed and the whole is placed under the bell jar of an air pump, it is observed that as the air is ex-

hausted from the receiver the globe settles down; when air is readmitted, however, the globe is again balanced by the weight. The globe with its greater volume displaces a greater volume of air than the weight, and by the law of buoyancy it must be buoyed up with a greater force.
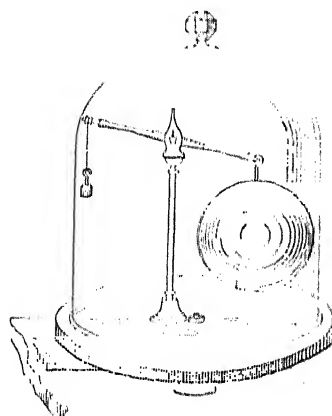
If a solid mass of brass is being weighed, using brass weights, the buoyant force of the air on both sides of the balance will be the same. But if the density of the weights is greater than that of the body weighed, the apparent weight of the body will be less than its true weight. When the apparent weight of a body is w, its true weight W may be found by the approximate formula,

$$W \qquad \qquad \left(\frac{1}{d} \quad \frac{1}{d_1}\right),$$

where δ is the density of air, d the average density of the object being weighed, and d_1 the density of the weights used.

204. Balloons. Balloons ascend in consequence of the buoyancy of the surrounding atmosphere. The gas within the envelope simply supplies the pressure to keep the balloon distended; in so far as it has weight it is a disadvantage. To find the supporting power of a balloon we must determine the weight of the balloon itself together with the enclosed gas and subtract this from the weight of an equal volume of atmospheric air. The difference is the portative force of the balloon.

As the balloon rises the pressure of the atmosphere decreases and the gas in the interior expands and completely fills the bal-

loon, and then as it expands still farther the excess escapes through an opening at the bottom.

EXPANSION OF GASES

205. Expansion of Gases. When a vessel containing gas is enlarged the gas expands, keeping the vessel full however great
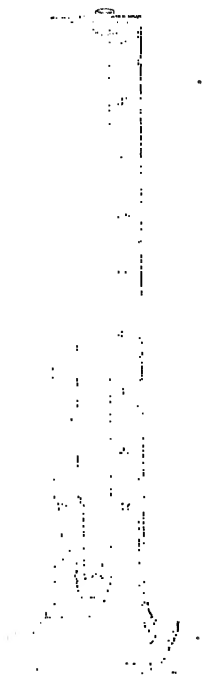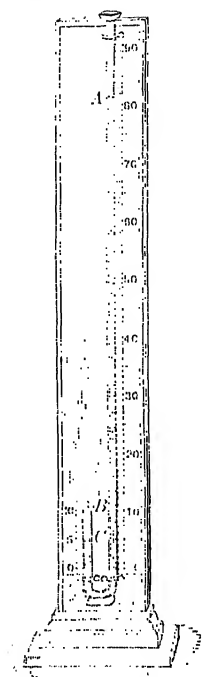
FIG. 105   FIG. 106

its volume may become, and at the same time the pressure of the gas diminishes.

If a small thin rubber bag containing a little air is closed and placed under the bell jar of an air pump, and the air exhausted from the space around the bag, the latter will be distended by the expansion of the enclosed air as the pressure upon it diminishes.

206. Boyle's Law. The exact way in which the pressure of a gas changes when its volume is varied was first investigated by

the English physicist Robert Boyle in 1662 and by Mariotte in France in 1679.

The form of apparatus used by Boyle is illustrated in figure 105. The short arm of the tube is closed and contains a mass of air separated from the outer air by the mercury in the bend of the tube. The enclosed air is at the same pressure as the outer air since the mercury stands at the same level in each branch. Mercury is now poured into the long arm of the tube until the enclosed air is compressed to one-half its original volume, as shown in figure 106. The height of the mercury in the long branch above that in the closed branch is then found to be just equal to the height of the barometric column. That is, the enclosed air is under a pressure of *two* atmospheres, one due to the external air pressure and the other due to the height of the mercury column.

If more mercury is added the air is still further compressed, and when the total pressure is three atmospheres, the mercury column having twice the barometric height, the air is found to be compressed to one-third of its original volume.

The law of compressibility of air, which is also found to be *approximately* true for all the more perfect gases may then be stated thus:

Boyle's Law. When the volume of a mass of gas is changed, keeping the temperature constant, the pressure varies inversely as the volume; or the product of the pressure by the volume remains constant.

That is, if a mass of gas has a volume v at a pressure p and if the volume is changed to v' while the temperature is kept constant, the pressure will become p' such that

$$pv = p'v' = \text{constant.} \qquad (1)$$

This constant is evidently proportional to the mass of gas used, for if the pressure is kept constant we must take twice the original volume in order to get double the mass of gas. We may, therefore, express Boyle's law by the equation,

$$pv = mk$$

or

$$\frac{pv}{m} = k \qquad (2)$$

where k is a constant which depends only on the kind of gas and its temperature.

Thus if we have a mass of gas m having pressure p and volume v, and another mass m' of *the same gas at the same temperature,* but with pressure p' and volume v', we have by (2)

$$\frac{pv}{m} \qquad \frac{p'v'}{m'} \tag{3}$$

Letting d represent the density of the gas, since $d = \dfrac{m}{v}$, we have from formula (3)

$$\frac{p}{d} = \frac{p'}{d'} = k; \tag{4}$$

that is, *the density of a gas is directly proportional to its pressure when the temperature is constant.* This is directly shown by Boyle's experiment, for with doubled pressure the volume is diminished to one-half and the density is consequently doubled.

To study the relation between pressure and volume for pressures less than one atmosphere, Mariotte used the apparatus shown in figure 107.

A long tube of glass closed at the upper end and plunged in a deep bath of mercury contains a small mass of air or other gas. The volume of the air or gas is given by graduations on the tube while its pressure is found by subtracting the height of the mercury column CD from the barometric height which measures the pressure of the external air. The volume and pressure are varied by raising or lowering the tube in the bath.
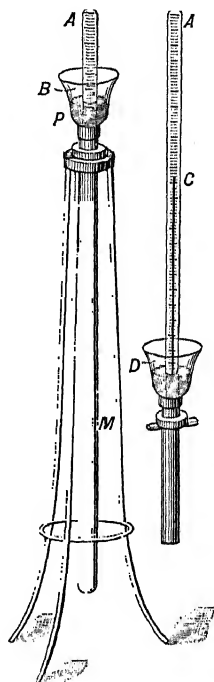
Fig. 107

207. Variations from Boyle's Law. Boyle's law is not *exactly* true in case of any actual gas.

The following table will indicate the degree of departure from the law, with increasing pressures, of some common gases:

VOLUME	PRESSURE IN METERS OF MERCURY			
	AIR	NITROGEN	CO₂	HYDROGEN
1	1.0000	1.0000	1.0000	1.0000
$\frac{1}{2}$	1.9978	1.9986	1.9824	2.0011
$\frac{1}{4}$	3.9874	3.9919	3.8973	4.0068
$\frac{1}{8}$	7.9156	7.9641	7.5193	8.0539
$\frac{1}{10}$	9.9161	9.9435	9.2262	10.0760
$\frac{1}{20}$	19.7198	19.7885	16.7054	20.2687

It will be noted that air and nitrogen are slightly more compressible than if they followed Boyle's law exactly, while hydrogen is rather less compressible; the departures from the law are, however, less than 1 per cent up to 10 atmospheres' pressure. Carbon dioxide shows marked increase in compressibility as the pressure increases and it approaches its point of condensation.

The French physicist, Amagat, has made an exhaustive study of the compressibilities of gases at different temperatures and up to pressures as great as 3000 atmospheres. More recently the American physicist, Bridgman, has carried the investigation up to the remarkably high pressure of 13,000 atmospheres. The results of both investigations show that as pressure is increased the product pv slightly diminishes at first, but when the pressure exceeds a certain amount, which depends on the gas and its temperature, the product pv steadily increases up to the highest pressures used.

The Dutch physicist, Van der Waals, has shown that the formula

$$\left(p + \frac{a}{v^2}\right)(v - b) = \text{constant},$$

in which a and b are small constants depending on the kind of gas, expresses quite exactly the relation of pressure to volume in gases at constant temperature for a far wider range of pressures than the simple formula of Boyle.

208. Measurement of Heights by Barometer. *The difference in pressure at two different heights in the atmosphere is equal to the weight of the unit column of air reaching from one level to the other.* If the average density of the air between the two levels were known then the height could easily be ascertained by dividing the difference in pressure by the average weight of unit volume of the air.

Let H represent the height in centimeters, P and p the two

pressures measured in grams per square centimeter, and d the average density in grams per cubic centimeter, then

$$Hd = P - p \quad \text{and} \quad H = \frac{P - p}{d}. \tag{1}$$

As the average density of the air between the two levels depends on pressure, temperature, and moisture, it is clear that the chief difficulty lies in determining this quantity.

Pumps and Pressure Gauges

209. Air Pump. Air pumps were first made by Otto von Guericke, of Magdeburg, in 1650. For rapid exhaustion when a vacuum of 0.1 mm. of mercury is sufficient, a very convenient pump is Gaede's rotary air pump, shown in figure 108 in which the cylinder A, mounted close to one side of a somewhat larger cylindrical cavity, is rapidly rotated by an electric motor and sweeps out the air from the crescent shaped space by means of two sliding vanes ss, which are carried in slots in A
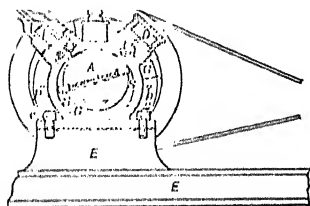
FIG. 108. Gaede rotary pump
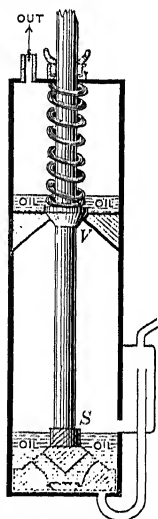
FIG. 109. Cylinder of Geryk air pump

and are pressed against the walls of the cavity by means of springs. In this way air is drawn in at C and forced out at D, finally escaping at J.

For higher exhaustion, pumps are used in which oil or mercury prevents leakage. In figure 109 the cylinder of the Geryk pump is shown in which a deep layer of oil covers the piston and valves so that no leakage of air back through the pump is possible.

When the piston is raised the air above is forced out through the valve V which is finally lifted by the shoulder S when the piston reaches the top, permitting the last bubbles of air to escape through the oil into the upper chamber, while at the same time oil flows down through the valve, filling the small space above the piston. In this way the air in the cylinder is *completely* expelled in each stroke.

Oil pumps for high exhaustions should never be operated without a drying tube to absorb all water vapor from the air before it reaches the pump, as moisture absorbed in the oil prevents the securing of a high vacuum.

A most effective pump of this type is one devised by Garde in which three cylinders, connected in series and mounted one above the other, form a single long cylinder and are operated with one piston rod. Air is drawn in at the bottom and forced successively through the three cylinders and escapes at the top. Only a small amount of oil is used and the presence of water vapor does not interfere with the action as it does in most oil pumps.

FIG. 110. Langmuir Condensation Pump

210. Langmuir Condensation Pump. A simple and convenient form of pump for producing high vacua has been invented by I. Langmuir, a form of which is shown in figure 110. Mercury, D, in the bottom of the cylinder, is heated by gas or electricity, causing it to evaporate. A stream of vapor flows upwards through E and is deflected by E downwards and against the water cooled sides A, where it immediately condenses, drops down and flows into D again through small openings. Molecules of mercury vapor flowing downwards from P strike against the molecules of the gas between E and A, crowding them downwards towards B, where they are drawn off. More gas is sucked in at C to displace that driven out until the vessel connected with C is very perfectly exhausted. The immediate condensation of the mercury vapor is essential to the operation of this pump, otherwise mercury vapor would flow upwards toward C and interfere with the downward suction. This pump is, therefore, called a condensation pump. Before this pump can operate, the pressure in it has to be reduced to a low value so that the molecules of the gas are far apart and can easily be driven downwards by the molecules of the mercury vapor. With higher pressures, where the
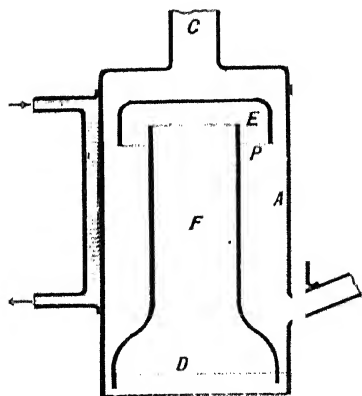
molecules of the gas are close together, the downward flow of the molecules of the mercury vapor is intercepted and exhaustion cannot proceed. The condensation pump is, therefore, operated in series with another type of pump which will give a sufficiently high exhaustion so that the condensation pump can operate. The Langmuir condensation pump combines simplicity with the power to produce an extremely high vacuum in a short time and is widely used commercially in the production of high vacua. Pressures as low as $\frac{1}{10,000,000}$ mm. of mercury may be obtained by means of this pump.

211. McLeod Gauge. For measuring the very low residual pressures in the vacua produced by air pumps, a device, shown in figure 111, and known as the *McLeod gauge* * is employed.

It is connected by the tube C with the exhausted vessel, so that the pressure in the bulb A is the residual pressure to be determined. The bulb D is raised causing the mercury to rise into A and C and compressing the air in A into the upper part of the narrow tube B. Suppose the air is thus compressed into one-thousandth part of the original volume $A + B$, the pressure in B will then be 1000 times the original pressure, while the pressure in the tube C is unchanged. The difference between the mercury levels in B and C will then measure the difference between the pressures, which in the case supposed is 999 times the pressure in C, so that 1 mm. difference in level corresponds to an original pressure of only 0.001 mm. of mercury.

FIG. 111. McLeod gauge

212. High Vacua. In obtaining the highest vacua chemical means also are employed. Sir Humphrey Davy was the first to use this method. Having put into the vessel to be exhausted some caustic potash and then filled it with carbonic acid gas, he pumped out the gas as far as possible, and, having sealed the vessel, left the residual gas to be absorbed by the caustic potash, and thus obtained a very good vacuum.

The principal difficulty in obtaining very high vacua is due to the escape of gases into the tube from the glass walls. These gases are actually dissolved in the glass. All ordinary glass contains considerable dissolved water vapor and carbon dioxide. When a certain degree of exhaustion is reached these gases escape into the tube as fast as they can be pumped out unless special means are employed to eliminate them. The difficulty is partly overcome by heating the glass tube during exhaustion to as high
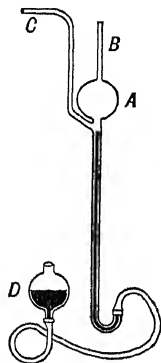
* Pronounced MacLoud.

a temperature as is safe without softening it. This drives out of the glass a considerable part of the dissolved gases which are then carried away by the pump. Metal electrodes, which may be sealed into the tube, also contain dissolved gases, but the metal electrodes can be heated to a very high temperature with-

out damaging them and so the dissolved gases are almost entirely driven out of them. But glass still retains dissolved gases even after heating. Chemical means are therefore employed to clean up these residual gases.
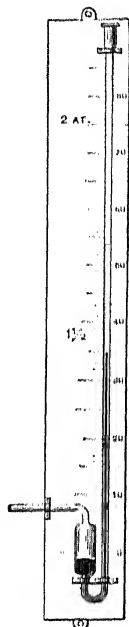
For instance, the tube may have a small piece of calcium sealed inside it which is vaporized by heating the tube after it is exhausted as far as possible by the pump. The residual gases combine with the calcium vapor and form products which readily deposit on the glass walls of the tube. The vapors of various metals can be used in this way to clean up the residual gases.

These high exhaustions are called by courtesy *vacua*, as they are the nearest approaches to an absolute vacuum that physicists have been able to make by the most refined methods known to science; and yet there is reason to believe that in every cubic inch of such a vacuum there are 400 million molecules of gas.

The very highest vacua obtained have been produced by means of vapor from a tungsten filament sealed into a tube and intensely heated by an electric current. At this high temperature, the tungsten vapor readily combines with the residual gases and the products of these reactions, consisting mostly of tungsten oxide and tungsten carbide, are deposited on the walls of the tube. The glass walls of the tube are chilled to a very low temperature by immersing the tube in liquid air. This so reduces the agitation of the molecules of the gases dissolved in the glass that they escape into the tube very slowly, making possible a remarkably high vacuum. By these means vacua higher than a million millionth of an atmosphere can be obtained.

213. Pressure Gauges. One of the simplest forms of pressure gauge is the *open manometer*. It consists of a bent tube containing mercury, one arm being open to the air and the other connected with the vessel in which the pressure is to be measured.

The difference between the pressure in the vessel and that of the atmosphere is measured by the height of one end of the mercury column above the other. If the difference in pressure to be measured is very small, it is often best to use water or even kerosene oil instead of mercury on account of their small densities.

214. Bourdon Spring Gauge. A device commonly used in steam gauges is the *Bourdon spring*, so called from its inventor. It consists of a tube of brass of elliptical section, bent into a nearly complete ring, the flatter sides of the tube forming the inner and outer sides of the ring. One end of the tube is closed and into the other the fluid under pressure is admitted by a pipe. This end of the tube is firmly fixed, while the closed end is free though connected with a pointer by levers and rack work or by a fine chain wrapped around a small spindle (Fig. 113) by which the motion is greatly amplified.
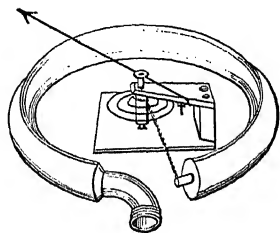
Fig. 113. Bourdon spring gauge

Suppose the pressure to increase, the flattened tube will spring a little and become more nearly circular in cross section, and in so doing it will slightly unbend as if to straighten out, causing the pointer to move over the scale. When this device is employed as a steam gauge the pipe leading to it is usually bent downward so that it fills with condensed water, preventing the hot steam from reaching the gauge.

215. Common Suction Pump. In this pump there are two valves opening upward, one in the piston and one at the bottom of the cylinder. As the piston is raised, its valve being shut, the atmospheric pressure forces water from the cistern to rise through the pipe and follow the piston, the lower valve opening and permitting this flow. As the piston descends the lower valve closes, preventing return to the cistern, and the valve in the piston opens allowing the water to pass through. Such a pump cannot raise water from a level more than about 34 ft. below the piston.

216. Force Pump. Water may be raised, however, to any desired height by the use of the force pump. In this pump the water is drawn into the cylinder as in the suction pump, but

the downward stroke of the solid piston forces the liquid in the cylinder out through the side tube into the rising pipe, which may be extended to any height. A valve in the side tube prevents flowing back, and an air chamber is provided which acts as a spring, the air yielding to sudden movements of the piston, which the water column on account of its great inertia could not do.

217. Siphon. If a bent tube is filled with a liquid and one end is introduced into a vessel of the liquid while the other end is open and held at a lower level than the surface, the liquid
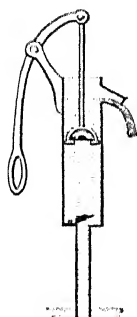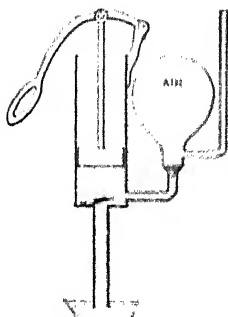
FIG. 114. Lift pump FIG. 115. Force pump FIG. 116. Siphon

will escape through the tube. Such an arrangement, known as a siphon, is represented in figure 116.

The upper surface of the liquid in the vessel and also the open end of the siphon are subject to the atmospheric pressure, but this is partly balanced on the short side by the column of liquid of height h, while on the other it is opposed by the longer column H. The pressure which is effective in causing the flow is, therefore, that of a column of liquid of height $H - h$. From this it appears that the velocity of liquid through a siphon would be the same as from an opening directly into the vessel at the level of the outer end of the siphon, if it were not for the loss due to friction in the pipe.

Clearly the liquid can only rise in the siphon to a height where it can be supported by the atmospheric pressure. water, there-

fore, cannot be lifted by a siphon more than 34 ft. above its level and mercury not more than 30 in.

218. The Centrifugal Pump. In the centrifugal pump liquid is admitted at the central part of a rotating wheel carrying veins or impellers, is caught between them and thrown outwards by centrifugal action, and escapes at the outlet provided for it even against a moderate pressure. More liquid is sucked in at the inlet to take the place of that discharged, and thus a steady flow is established. This pump is useful for lifting large volumes of water short distances. Similar pumps called *blowers* are used for handling large volumes of air for ventilation purposes or for establishing forced drafts in furnaces, etc.
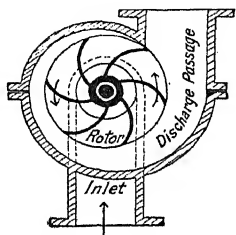
FIG. 117. Centrifugal pump

PROBLEMS

1. How high would the atmosphere have to be to cause the barometer to stand 76 cm. high, if its density was the same throughout as at the earth's surface, taking this density as 0.0012 gm. per c.c.?

2. How much higher will a barometer stand at the base of a mountain than at a station 1000 meters higher, taking the average density of air between the stations as 0.0012?

3. The air chamber of a force pump contains at the start 600 cu. in. of air at pressure 75 cms. of mercury. What volume will the air occupy while water is being forced to a height of 150 ft. above the pump?

4. How deep must a pond be that an air bubble on reaching the surface may have twice the volume that it had at the bottom? Suppose the barometric pressure at the surface to be 75 cm. of mercury.

5. How deep must a pond be when a bubble having a volume of 12 c.c. at the bottom has a volume of 30 c.c. as it reaches the surface? Barometer reading 75 at the surface.

6. A barometer on top of a tower stands at 75.20, at the bottom it stands at 75.40. How high is the tower if the average density of the air between the top and bottom is 0.0012 gm. per c.c.?

7. A barometer having a little air in the top of the tube stands at 72; but if the level of the mercury is raised so that the air space is half as great as before, it stands at 70. What is the correct barometric height?

8. If the tube in the apparatus shown in figure 107 ...ains 100 c.c. of air, and the mercury stands in the tube 15 cm. above t..... vel in the outer vessel, while the barometer stands at 75, find what would be the volume of the enclosed air if it were at atmospheric pressure ... it will the volume of the enclosed air become when the tube is raised ....iently to make the mercury stand 20 cm. high inside the tube?

9. A glass bottle containing 100 c.c. of air at atmospheric pressure floats at the surface of a pond with its open mouth downward. The bottle weighs 130 gms. and the density of the glass is 2.6. If the barometric pressure is 75 cm. of mercury, how deep below the surface must the bottle be pushed that it may just float in equilibrium, neither tending to rise nor sink? Neglect the weight of the enclosed air. Will the equilibrium be stable or unstable and why?

10. What force must be exerted on the piston of a force pump 3 in. i diameter to raise water 100 ft.?

PART II. FLUIDS IN MOTION

219. Steady Flow. When a fluid is in motion if the pre...ure, velocity and direction of flow remain unchanged at every point in a certain region, the motion there is said to be steady. A line drawn in the fluid so that at every point it is in the direction of the flow at that point is called a *stream line*.

220. Continuity. In case of steady flow as much fluid must flow into any region as flows out of it in the same time.
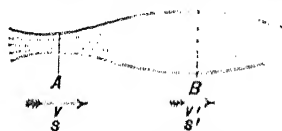
FIG. 118

Let the figure represent either an open channel or a pipe conveying water. The total volume of water crossing the section of *A* per second will be vs cu. ft. per second if the velocity is v ft. per second and the cross section of the stream at that point is s sq. ft. If d represents the density at *A*, or the number of pounds mass per cubic foot, then vsd is the mass of water crossing *A* per second and similarly $v's'd'$ is the corresponding mass crossing *B* in the same time, and therefore $vsd = v's'd'$. This equation holds for the steady flow of any fluid whether gas or liquid. But for liquids since the density does not appreciably change during the flow, we may take $d = d'$ and so

$$vs = v's'$$

or the velocity is inversely as the cross section of the stream. If at a narrow place in a stream the velocity is not correspondingly great, we may be sure that the stream is deep at that point. The extremely small cross section of a stream at the edge of a dam is due to its great velocity at that point.

221. Momentum of Liquid Stream. When a liquid is in motion each moving particle has momentum and kinetic energy. When a jet escapes through an opening in the side of a vessel the pressure which gives the jet its forward momentum acts at the same time as a reaction pressing the vessel in the opposite direction. If the orifice is free to move backward it will do so, as in case of the device known

FIG. 119.   Barker's mill

as Barker's mill shown in the figure.   In case of the end of a hose the rush of water around a curve will by its centrifugal force tend to straighten the hose.   If the end is free it will very probably swing over too far, in consequence of its inertia, when it will be flung back again, thus thrashing to and fro.
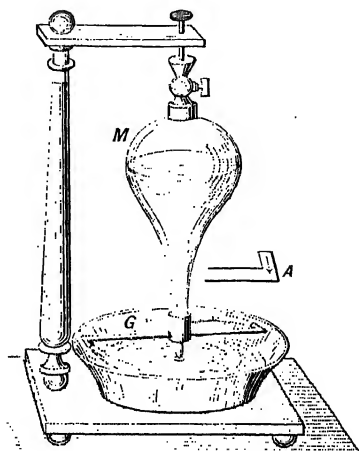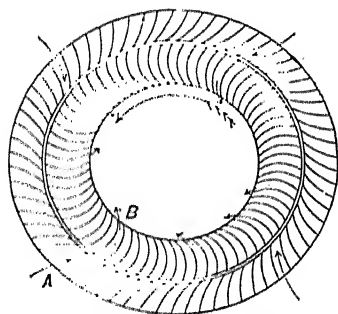
FIG. 120.   Turbine water wheel diagram

222. Turbine Water Wheels. The centrifugal force of a stream as it moves by curved guides is made use of as a means of obtaining power in turbine water wheels. Such a wheel is shown in section in the diagram. The water flows inward toward the wheel through the fixed guides, which cause it to enter in the proper direction, and then driving the wheel forward and sweeping by the wheel guides *BB*, it escapes at the center of

the wheel. The guides *AA* may be made adjustable so as to regulate the flow of water. The entering water from the flume is conducted to the turbine by a pipe which is kept constantly full, thus giving the advantage of its pressure. The turbine may be set at the lowest level so that the water escapes directly into the tail race, or it may be set higher if the water escaping from the wheel enters a closed draft pipe which leads down to the tail water.

The sinking of the water in this draft pipe produces a suction which increases the efficiency of the wheel.

223. Efficiency of Water Wheels. When water flows from one level down to another it loses potential energy. That proportion of the potential energy lost by the water which is transformed into useful work in a water wheel is called its *efficiency*. It is clear that to be efficient a wheel must as far as possible let the water down from the higher to the lower level without churning, and the water escaping at the bottom should have little velocity, its energy having been expended in useful work.

224. Various Water Wheels. The old fashioned overshot wheel, taking water from the upper level and lowering it to the bottom of the fall, uses the whole energy of the fall, but its size and weight cause great frictional loss.

Where a small supply of water at high pressure is available, some form of jet wheel is often best. Here the wheel is driven at high speed by the force of a jet escaping against cups set around the periphery of the wheel.

225. Hydraulic Ram. The hydraulic ram is an appliance by which a small quantity of water may be raised a considerable height by using a small fall in a stream. The water is conducted to the ram through a straight, smooth, inclined pipe offering little resistance to the flow. At *C* is a valve opening *downward* through which the water at first escapes; but as its speed increases, it catches the valve in its rush and shuts it. This sudden stoppage of the stream causes a great pressure at this end of the pipe in consequence of the forward momentum of the stream, and the valve *d* which opens *upward* is forced open and some water driven into the pipe *e*. The valve *d* then closes and prevents any return of water from *e*. But with the sudden stoppage of the stream the valve *C* if properly weighted rebounds and

opens again, the stream again escapes at C with increasing velocity until the valve is again caught and closed, when water is again driven through the valve d by the hammer-like blow of the column of water in A. The action is thus kept up indefinitely, water being gradually forced up the pipe e until it may reach
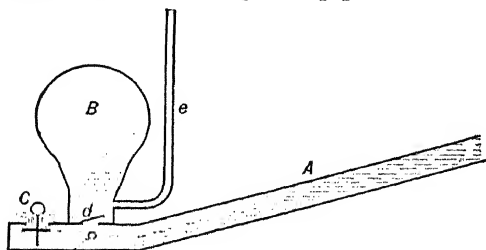
FIG. 121.   Hydraulic ram

many times the height through which the stream falls. The air chamber B is essential to the action of the ram as it presents an elastic cushion with but little inertia, enabling the valve d to yield instantly. At f there is a minute opening, the *air sniff*, through which, in the recoil of the water, air is drawn in, maintaining the supply in the air chamber. *If a hydraulic ram were perfectly efficient*, it would raise one-tenth of the amount of water flowing into it through ten times the height of the fall or one-half the water twice the height of the fall. But in practice the efficiency of a good ram is about 50 per cent.

FIG. 122

Rams are now made in which the supply pipe is as much as 4 ft. in diameter. In these rams the valve which arrests the flow is moved by a piston operated by water from a small branch of the main pipe.
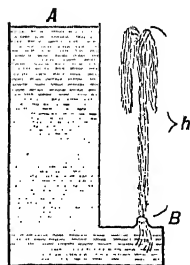
226. Velocity of a Jet. While a liquid is escaping from a vessel through an opening which is small compared with the upper surface of the liquid, no *change* takes place within the vessel except the gradual lowering of the surface or disappearance of liquid from the top, while a corresponding mass appears outside in the escaping jet. If no energy is lost in friction or viscosity, the energy of a mass escaping at B must be the same

as the energy of an equal mass at A. But since the potential energy at B due to gravity is less than at A the kinetic energy at B must be correspondingly greater; that is, it must be great enough to cause the escaping mass to rise from B to A when the jet is directed upward.

If h is the height of A above B we have

The difference between the potential energy of a mass m at A and $B = mgh$ ergs.

The kinetic energy of mass m escaping at B with velocity $v = \frac{1}{2}mv^2$ ergs.

Therefore

$$mgh = \frac{1}{2}mv^2$$

and

$$2gh. \tag{1}$$

This velocity is the same as that which a freely falling body would acquire in a distance h, a conclusion known as *Torricelli's theorem*.

Torricelli's Theorem. The velocity of an escaping is equal to the velocity which a body will acquire in falling from the level of the upper surface to that of the opening.

The density of the liquid and direction of the jet do not affect its velocity.

When the pressure alone is known, the height of the liquid required to produce the given pressure may be calculated and then used in the above formula. Thus the pressure on the level of B is $p = hdg$ in *dynes per cm²*; using this to eliminate h from equation (1) we obtain

$$v^2 = \frac{2p}{d}. \tag{2}$$
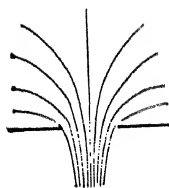
227. Vena Contracta. The liquid as it approaches the opening moves in from all sides along *stream lines* like those shown in the diagram. Liquid coming from each side has a certain momentum toward the axis of the jet, hence the jet narrows and does not become cylindrical until just after it has left the orifice. A short cylindrical neck of the size of the opening is found to increase the quantity escaping per second, and if the neck is somewhat flared out the flow is still greater.

228. Efflux of Gases. The velocity with which a gas escapes through a small opening when the difference between the pressures on the two sides of the opening is p is also determined by Torricelli's theorem.

$$v = \sqrt{\frac{2p}{d}}.$$

Since for a given pressure the velocity of efflux is inversely proportional to the square root of the density of the gas, the densities of gases may be compared by observing the times in which measured quantities escape through a small opening.

229. Energy Due to Pressure. *When a liquid is forced into a vessel against pressure, the work done is equal to the product of the pressure by the volume of the liquid which is introduced.* This expenditure of work is not wasted in friction, but exists as energy in the mass, ready to be transformed into energy of motion if an opening allows the mass to escape. The amount of this energy, since the volume of the mass m equals $\dfrac{m}{d}$, is

$$E = \frac{pm}{d}.$$

230. Energy Equation. Consider a small mass of liquid at a in the vessel shown in the diagram; it is in equilibrium, and may be moved without offering any resistance from a up to the surface. Clearly there is no change in its total potential energy as it is moved from one part of the vessel to another. At the top its gravitation potential energy referred to the earth is a maximum, but then it has no energy due to pressure, while at a its gravitation energy is less but its pressure energy is correspondingly greater. If h represents the height of the mass m above some fixed plane, say the surface of the earth, its gravitation potential energy referred to that plane is mgh. We have seen that its pressure energy is $\dfrac{mp}{d}$; and if the mass is in motion it will have kinetic energy $\frac{1}{2}mv^2$, and its total energy may be written

$$\frac{mp}{d} + mgh + \tfrac{1}{2}mv^2 = \text{energy of mass } m.$$

If the stream is flowing in conduits or channels without doing work, the energy of the mass will remain constant except as it is wasted in internal or external friction. The fact that in steady irrotational motion of a friction-less fluid, the expression $\dfrac{mp}{d} + mgh + \frac{1}{2}mv^2$ remains constant for a little mass m as it moves along, is known as *Bernoulli's Principle.*

As an illustration of the above equation, conceive the vessel A in the figure to be kept filled to a constant level while the liquid is flowing out

freely through the pipe D, and follow the changes in the mass m. As it moves downward h grows less and so its gravitation energy diminishes while the pressure energy increases, the kinetic energy being scarcely changed, but as it approaches the opening B its velocity increases, and consequently more of its energy is kinetic and less due to pressure than at the same level farther in the vessel, so the pressure at B must then be less than at b. When it reaches C its velocity will be less, and consequently the pressure there will be greater than at B. Throughout D the cross section, and consequently the velocity, is constant, and since it is all at the same level the pressure must be constant except as influenced by friction in the pipe.
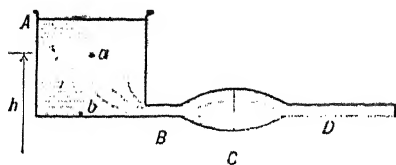
FIG. 124

231. Friction in Pipes. When water escapes from a reservoir through a horizontal pipe of uniform section, as ab in figure 125, the velocity will be the same at all points in the pipe, and *if there is no friction* the pressure will be constant throughout the length of the pipe and equal to the atmospheric pressure at the end b. In that case the water will not rise in any of the gauge tubes shown. In practice, however, there is always some friction in a pipe, and, therefore, a constant expenditure of energy. But the energy equation is

$$E = \frac{mp}{d} + mgh + \tfrac{1}{2}mv^2,$$

and if the pipe is level and cylindrical h and v cannot change, consequently if there is any decrease in E there must be an equal decrease in the first

term $\frac{mp}{d}$, and, therefore, a *fall in pressure*. If the friction is uniform throughout the pipe the pressure will decrease uniformly, becoming equal to the atmospheric pressure at the opening b. The height h (see figure) which determines the pressure at b when the opening is stopped up so that there is no flow,

FIG. 125

is called the *pressure head*. When b is open the head required to produce the observed velocity of escape reckoned from the law $v = \sqrt{2gh}$ is called the *velocity head*. In the above case it is h', and the remaining head $(h - h')$ is spent in overcoming friction.
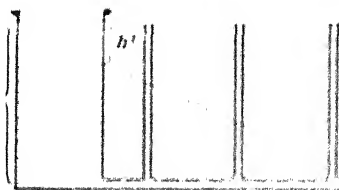
The loss of pressure when water is flowing in pipes is a fact that has to be constantly taken into account in practice. The friction and consequent loss of pressure, increase with the velocity of flow.

232. Pressure Varies with Velocity. The fact just demonstrated that in a horizontal pipe of variable section *the pressure will be greatest when the cross section is the greatest and velocity least* is so interesting and important that it merits a brief examination from another point of view.

Consider a little mass of liquid at A (Fig. 126) where its velocity is clearly diminishing. It is under pressure on all sides, but since its velocity is diminishing the pressure backward on its forward side must be greater than the pressure on its left which urges it forward. When the mass reaches B, however, its

FIG. 126

velocity is increasing, hence the pressure behind it which urges it forward must be greater than the pressure in front which is opposite to its motion; the point of slowest motion must therefore be a point of maximum pressure.

233. Aspirating Pumps. The principle just established is made use of in aspirators for exhausting air. Such an instrument

FIG. 127. Chapman exhaust pump

is shown in figure 127. It provides a narrow channel through which water flows with great velocity, the stream widening out and moving slower before it reaches the atmospheric pressure at the open end. The pressure at the narrowest point must then be very much less than that of the atmosphere, and air is accordingly drawn in through the side tube C and carried out at B by the rush of water.

234. Ball on Jet, etc. A jet of water or even of air may support in stable equilibrium a light ball. The explanation is that a slight shifting of the ball, say to the right, would cause the main stream to rush on the left, the velocity of flow would be greatest there, and therefore, the pressure less than on the right, and so the ball would be pressed back again.

A card with a pin through it and laid over the open end of a spool cannot be blown off by blowing through the spool because the velocity of the air stream as it spreads out under the card is least at the outer edge where it comes to the atmospheric pres-

sure, the pressure nearer the center where the velocity is greater will, therefore, be less than that of the atmosphere and the card will accordingly be pressed against the end of the spool.

Similarly a ball will be held in a cup by a jet escaping around it, as in the ball nozzle used for fire hose.

If a stream of air is directed between two sheets of paper they are drawn together. So also leaves and other light objects are drawn *toward* a moving train as it passes.

PROBLEMS

1. Find the velocity of the stream of water in a pipe having a cross section of 3 sq. in. and discharging 450 cu. ft. of water per hour.

2. How much water will escape per minute from a 2 in. hole in the side of a water tower 50 ft. high?

3. With what velocity will water spurt out of a hole in a boiler in which the pressure is 80 lbs. to the sq. in., in addition to atmospheric pressure?

4. The stream of water below a certain dam has a cross section of 10 sq. ft. and a velocity of 5 ft. per sec. Find the horse power available if the dam is 15 ft. high.

5. What horse power would be obtained from a 20 ft. fall by a turbine wheel of 80 per cent efficiency, when the flow is 300 cu. ft. per minute?

6. If the water escaping from a turbine water wheel which uses the water from a 10-ft. fall has a velocity of 6 ft. per sec., what is the greatest possible efficiency of the wheel?

7. If the efficiency of a hydraulic ram is 50 per cent, how much water per day will it raise to a tank at a height of 100 ft. above the ram, when the supply pipe has a fall of 8 ft. and discharges 1 gallon per minute?

8. While water is flowing with a velocity of 2.2 ft. per sec. in a pipe 1 in. in diameter the pressure drops off from 70 to 10 lbs. per sq. in. in a length of 500 ft. Find the energy in foot-pounds spent in overcoming friction per cu. ft. of water.

9. Find the horse power spent in friction in the 500 ft. length of pipe specified in problem 8.

10. Derive formulas (1) and (2) of § 226 for the velocity of an escaping jet, from the energy equation of § 230.

11. A horizontal water pipe of 1 sq. in. cross section widens out to 3 sq. in. in section. If the velocity is 5 ft. per sec. in the narrower pipe and the pressure 5 lbs. to the sq. in., what will be the pressure in the adjoining part of the wider pipe?

Ans. 5.14 lbs. per sq. in.

The pressure given is *gauge pressure*, or the excess above that of the atmosphere. The total or absolute pressure is $5 + 14.7 = 19.7$ lbs. per sq. in.

12. The nozzle of a fire hose has an opening 2 in. in diameter, while the pipe just back of it is 3 in. in diameter. Find the pressure just back of the nozzle when it can throw a jet 60 ft. vertically upward.

Ans. 20.9 lbs. per sq. in.

NOTE: At the opening of the nozzle the pressure is that of the atmosphere, or 14.7 lbs. per sq. in. absolute, while the velocity is found from the height to which the water is thrown. Use energy equation of § 230.

PROPERTIES OF MATTER

AND ITS

INTERNAL FORCES

STRUCTURE

235. Density. On comparing a block of wood or aluminum with an equal weight of lead or gold, it is clear that substances differ greatly in the quantity of matter concentrated in a given volume. *The mass of any substance contained in unit volume is known as its density.*

DENSITIES OF SOME SUBSTANCES IN GRAMS PER CUBIC CENTIMETER

Solids		Liquids		Gases at 0° C. and 1 Atm.	
Aluminum..	2.7	Mercury......	13.596	Air	0.001293
Iron.......	7.2 7.8	Sea water.....	1.026	Oxygen	0.001430
Tin........	7.3	Water at 4° C.	1.00	Nitrogen	0.001256
Copper.....	8.8	Alcohol.......	0.8	Hydrogen	0.00008988
Lead.......	11.4	Ether........	0.72		
Gold.......	19.3				
Silver......	10.5				
Platinum...	20.5 22.0				
Tungsten...	18.6 19				
Brass......	8.3 8.6				
Glass......	2.5 3.5				
Oak wood..	0.84				

236. Molecular Forces. When a lead bullet is divided by a clean cut, if the two halves are pressed together they will cling with considerable force. This force is of the same nature as that which originally held the two parts together, but is far smaller because of poor contact. This force is known as cohesion, or when the attraction is between different substances it is known as adhesion. All substances attract each other in this way in greater or less degree. A drop of water is held together by cohesion, but it clings to a glass rod by adhesion. Cohesion and adhesion are of exactly the same nature, the sub-

division being merely one of convenience and otherwise of not much importance.

237. Molecular Theory. All matter is conceived as made up of separate *molecules* which are the smallest portions of the substances that can exist in a free state, as in gas or vapor. The molecules of any particular substance are all alike, and, in substances not at the absolute zero of temperature, are in more or less active motion or vibration, the energy of vibration depending on the temperature. In solids the vibrating molecules are held by their mutual attractions in such a way that they cannot move far away from their mean relative positions. In liquids the phenomena of diffusion, and the Brownian movement (§ 282), show that molecules move about in the mass, and are not held in fixed positions relative to each other, though the force of cohesion may be very great. In gases or vapors there is the greatest freedom of motion of the molecules, and their average distance apart is much greater than in liquids or solids, while there is scarcely any cohesion.

It is supposed that any two molecules of matter attract each other, according to the Newtonian law of gravitation, with a force varying inversely as the square of the distance between them for all considerable distances, but when very near each other the force of attraction varies with the distance according to some unknown law, giving rise to the phenomena of cohesion and adhesion, until the molecules come into what is called contact, when a force of repulsion opposes nearer approach.

The experiments of Quincke indicate that molecules must be less than 5×10^{-6} cm. apart in order that the cohesive force may be perceptible.

The idea that matter is molecular in its structure is supported by a great variety of evidence found especially in the phenomena of heat, gases, and radiation, as well as in chemical phenomena.

238. Molecular Structure. The molecules themselves are made up of one or more *atoms*. In most gases such as hydrogen or nitrogen, each molecule consists of a pair of two atoms, held together by their mutual attractions. There are about eighty-seven different kinds of atoms distinguished from each other by their different chemical properties. A substance whose molecules

are made up of only one kind of atom is called an *element*.   Hydrogen, whose molecules consist of two hydrogen atoms each, is an element.   So also are gold, silver, and iron.   Water, on the other hand, is made up of molecules consisting of two hydrogen atoms and one of oxygen, and is called a *compound*.   So also is common
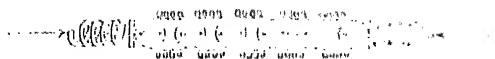
FIG. 128

salt whose molecules are made up of one atom of sodium and one of chlorine.

Even the atom, itself, can be subdivided.   Each kind of atom consists of a central nucleus surrounded by a definite arrangement of a certain number of *electrons*.   The *electron* is the smallest known particle of matter, having a mass equal to about $\frac{1}{1840}$ of that of the hydrogen atom, and as we shall see later, it has a definite electric charge from which it cannot be separated (§ 560).   It is thus an atom of electricity.   The electron is a common constituent of all elements and hence of all matter.

239. Molecular Equilibrium.   The molecules of a substance may be regarded as in a state of equilibrium under three forces; external pressure, cohesive force, and an internal pressure due to the rebounding of adjoining molecules against each other as they vibrate to and fro.   This latter force may be considered to balance the other two.

We may form a conception of these forces by the following model. Imagine a row of small rubber balls drawn together by springs stretched between them (Fig. 128).   Let two outer springs also press them together, and let the balls be thought of as rapidly vibrating to and fro, rebounding against each other, and so keeping a greater distance apart than if they were at rest.

The force of the outer springs represents the external pressure and that of the springs joining the balls **together** the attractive force,* and these are balanced by the repulsion due to the impacts of the balls against each other.

* The springs representing the attractive force should be conceived as exerting less force the more they are stretched, for cohesive force diminishes as particles separate.

In solids and liquids the force of attraction between atoms and molecules is that which chiefly balances the internal repulsion, the external pressure being usually quite insignificant in comparison.

But in gases the case is different. In consequence of the great average distance between the molecules, the cohesion is so insignificant that the external pressure alone may be said to balance the internal pressure due to the motions of the molecules.

This theory of gaseous pressure is more fully discussed in § 279 *et seq.*

240. Structure. When the properties of any one portion of a mass are exactly like those of any other portion, the mass is said to be homogeneous. Whether a substance is called homogeneous or not depends on the point of view. One part of a brick wall is just like another part, and so it may be said to be homogeneous; but if we compare minute parts we find in some spots brick and others mortar and so there is a limit to its homogeneity. So water is regarded as homogeneous unless we are dealing with portions so small that the molecular structure is significant.

If the various physical properties of a substance are the same in all directions throughout its mass, it is said to be *isotropic*. Water, glass, and mercury are isotropic. Most crystalline substances are not isotropic, and may be called *anisotropic*.

241. Crystals. In solids which pass slowly into the solid state, either directly from vapor or as the result of the slow cooling of a fused mass or of separation from a solution, there are often formed masses called crystals which have regular and distinctive forms and are bounded by plane faces.

The crystallization begins at certain isolated points and the minute crystals gradually grow in size, until they may meet and form a solid agglomeration.

The study of the fundamental crystal forms has led mineralogists to divide them into six classes or systems.

Some idea of the cause of the formation of crystals may be obtained by considering the forms which may be built up of shot when placed together so that each shall touch as many others as possible. Suppose such a pyramid as that represented in figure 129, where one layer is incomplete; if we think of it as a growing crystal in which the balls represent the atoms and suppose it im-

mersed in a medium in which there are free atoms surrounding it, there will clearly be a tendency for these to fill out the incomplete surface, for an atom will touch more neighbors when placed along the incomplete edge than anywhere else, and so may be conceived to be more powerfully attracted into that position.
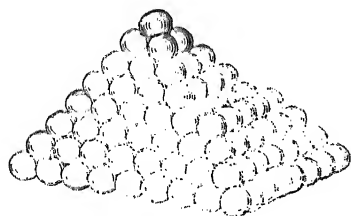
FIG. 129

In consequence a figure bounded by plane surfaces would result.

The piling of balls would give the crystal forms characteristic of the first or regular system, but to explain the variety of crystal groups it is necessary to suppose that the atoms themselves have properties different in one direction from what they have in another, and that when built up into crystal forms they are all similarly oriented or directed.

The exact atomic arrangement in many varieties of crystals has been found by the use of X-rays as described in § 979.

ELASTICITY AND VISCOSITY

242. Stress and Strain. When a portion of matter is acted on by forces tending to change its size or shape it is said to be under *stress*, and the accompanying distortion or change in volume is called the *strain*.

A stress tending to stretch any portion of matter is called a *tension*, while a stress tending to shorten it is called a *pressure*.

Stress is measured by force per unit surface, as in pounds per square inch, or in grams or dynes per square centimeter.

243. Strain Ellipsoid. When a body is strained, a small spherical portion of it is in general distorted into an ellipsoid and the axes of the ellipsoid are the three principal directions of strain at that point.

When the strain is the same everywhere throughout a body, as in case of a stretched wire, it is said to be *homogeneous*. In such a case the strain ellipsoids are all alike and similarly situated, as shown in figure 130.

When a fluid is compressed the strain is homogeneous and the ellipsoids are spheres slightly smaller than in the unstrained state.

The distribution of strain in a bent beam is shown by the ellipsoids in figure 131. The strain in this case is not homogeneous and there is a surface of no strain indicated by the dotted line.

244. Resistance to Strain. A body is said to be *elastic* if after having been strained it springs back to its original form when the stress is removed. If the stress is the same for a given amount of strain, whether the strain is increasing or diminishing, the body is said to be *perfectly* elastic.

When strained beyond a certain point called *the limit of elasticity*, substances yield permanently and do not return to the original state when the straining forces are removed. In this case there may be a great internal stress while the body is being
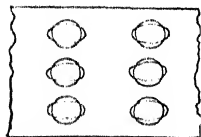
Fig. 130
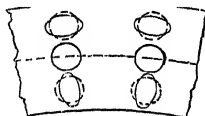
Fig. 131

strained, but on a very slight diminution of strain the stress entirely disappears. Putty, wet clay, and lead all exhibit this permanent distortion under comparatively small forces and even when the strain is small; while india-rubber is remarkable for the great strain which it can experience without passing its elastic limit. It is said to have a *wide* limit of elasticity.

Even within the limits of elasticity most substances show a time lag in returning to their original state after having been strained. Thus when a steel wire is firmly clamped at its upper end, if the lower end is twisted through an arc well within its limit of elasticity, the wire when set free returns at once nearly to its original position, but creeps very slowly back through the remaining distance. This lag is found in metals and in glass but quartz fibers are remarkably free from it.

245. Hooke's Law. *In small strains of elastic bodies the stress is proportional to the strain.* This is known as Hooke's law, having been enunciated by him in 1676. According to this law,

a long spring when stretched 2 cm. will exert twice the force that it would if stretched 1 cm., and the tension required to stretch a spring a small distance is equal to the pressure when the spring is compressed an equal amount.

Careful experiment, however, shows that the law is not exactly true. Most substances offer slightly more resistance to a given small compression than to an equal extension.

An illustration of this law is afforded by the ordinary spring balance in which equal divisions of the scale correspond to equal increments of weight. In this case the elongation or compression of the helical spring may be relatively very great, yet because of its shape the distortion or strain of any little portion is extremely minute and Hooke's law holds very nearly true.

246. Elasticity. In elastic bodies the elasticity is measured by the ratio of the stress to the corresponding strain.

$$\text{Elasticity} = \frac{\text{stress}}{\text{strain}}.$$

In bodies which are homogeneous and isotropic there are two principal kinds of elasticity, that in virtue of which the body resists change of volume and that resisting change of shape.

The first is called volume elasticity and the second rigidity. Volume elasticity is possessed by all bodies, fluids as well as solids, but rigidity is a characteristic of solids.

In some strains both of these elasticities are involved; for instance, when a wire is stretched there is a sidewise contraction as well as an elongation, so that the resistance to stretching depends on both the rigidity and volume elasticity of the substance. The elasticity of stretching or compression is so important in engineering that it has received a special name and is known as Young's modulus.

247. Volume Elasticity. When a body is so strained that every little cubical portion is compressed into a smaller cube the corresponding stress must be a pressure equal in all directions, provided the substance is isotropic or equally compressible in every direction.

This kind of stress is called hydrostatic pressure because it is the only kind of stress that can exist in fluids at rest.

The *volume elasticity* or *bulk modulus* of a substance is the ratio of the increase in pressure to the corresponding compression per unit volume.

Thus this elasticity will be represented by

$$E = \text{Volume elasticity} = \frac{\text{pressure increase}}{\text{change in unit volume}} = \frac{p}{\dfrac{v}{V}} = V\frac{p}{v},$$

where p is the increase of pressure caus-
ing a contraction v in a total volume V.

The volume elasticity of a solid may
be found by subjecting a long bar of the
substance to hydrostatic pressure in a
strong tube having thick glass windows
through which its change in length may
be observed by fixed microscopes.

248. Compressibility of Liquids.
Liquids are, as a rule, somewhat more
compressible than solids, but on the
other hand so great is their resistance to
compression that for most practical
purposes they may be treated as if in-
compressible.

The compressibility of a liquid may
be measured by the apparatus shown in
figure 132, known as Oersted's *piezom-
eter*. In this instrument the liquid to
be tested is contained in a bulb of glass
terminating in a long narrow tube of uni-
form diameter, open at the end and care-
fully graduated. This bulb A is sur-
rounded by water in a stout cylindrical

Fig. 132. Oersted's
piezometer

vessel of glass and subjected to pressure by means of a piston
forced in by a screw. A globule of mercury in the narrow tube
separates the liquid in the bulb from the surrounding water.
From the number of scale divisions through which the mercury
moves down toward the bulb as pressure is applied, the *apparent*
compressibility of the contained liquid is determined, the relation
between the volume of the bulb and the volume contained in one

division of the capillary tube having been previously ascertained. Although the pressure is the same on the outside of the bulb as on the inside, its volume *diminishes* in consequence of the compression of the glass of which it is made, so that the experiment gives the *difference* between the compressibility of the liquid and that of the glass bulb.

A thermometer gives the temperature of the liquid examined, and the pressure may be determined from the amount of compression observed in a tube M containing air and placed open end downward in the cylinder.

249. Elasticity of Gases. In case of a gas it is necessary to distinguish between its elasticity when the temperature is kept constant during the compression, and its elasticity when compressed so suddenly that there is no time for the flow of heat to take place. The first is called isothermal elasticity and the second adiabatic elasticity; the latter is always greater, being, in case of air, oxygen, hydrogen, and nitrogen, about 1.40 times as great as the isothermal elasticity.

The isothermal elasticity of a gas may be calculated from Boyle's law. Suppose the pressure is increased from p to p', the decrease in volume will be $v - v'$ and we have

$$E = \frac{p' - p}{\ } \qquad (1)$$

But by Boyle's law $pv = p'v'$, therefore

$$\frac{p}{p'} = \frac{v'}{v}$$

and

$$\frac{p' - p}{p'} = \frac{v - v'}{v}.$$

Substituting in (1) we find

$$E = p'.$$

But the difference between p and p' is supposed extremely small, so that *for gases kept at constant temperature the volume elasticity is equal to the pressure.*

VOLUME ELASTICITY AND COMPRESSIBILITY

| SUBSTANCE | TEMP. | COMPRESSIBILITY IN MILLIONTHS OF VOLUME PER ATMOSPHERE | VOLUME ELASTICITY | |
			DYNES PER SQ. CM.	LBS. PER SQ. INCH
Steel............	...	0.55	188.00×10^{10}	27.00×10^6
Glass............	...	2.44	41.00×10^{10}	6.00×10^6
Mercury.........	0°	3.00	33.00×10^{10}	4.8×10^6
Glycerin.........	20°	25.00	4.0×10^{10}	0.58×10^6
Water...........	20°	46.00	2.2×10^{10}	0.32×10^6
Ether...........	20°	191.00	0.52×10^{10}	0.07×10^6
Air { At normal pressure }		1,000,000.00	1.00×10^6	14.7

250. Rigidity. If a cylindrical rod or wire is twisted about its axis without change of length it may be imagined divided into sections of equal thickness, in each of which there has been no change in volume but simply a distortion of the little elements of which it may be conceived as made up.

FIG. 133

Take such a little block as that represented in figure 133. If the base CD is firmly fixed, a force F applied to the upper surface will strain it into the position $A'B'$, just as a thick book lying on a table may be pushed out of shape by force applied to the upper cover. The strain in this case is a pure distortion without any change in volume and is called a *shear*, and the forces bringing it about constitute a *shearing stress*. The strain is measured by the ratio of the displacement AA' or x to the height h, while the stress is the force applied per unit area; or if S is the area of the upper surface of the block the stress is $\frac{F}{S}$. The rigidity n, or elastic resistance to distortion, may therefore be expressed thus:
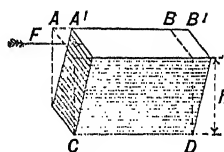
$$\text{Rigidity} = n = \frac{\dfrac{F}{S} = \text{stress}}{\dfrac{x}{h} = \text{strain}} = \frac{Fh}{Sx}. \tag{1}$$

In case of a wire clamped at one end and twisted at the other, it may be mathematically demonstrated that the moment of the force of torsion T is expressed by the formula

$$T = n \frac{\pi r^4 a}{2l},$$

where n is the coefficient of rigidity of the substance of which the wire is made, r is the radius of its cross section, and l its length, while a is the angle (in radians) through which it is twisted.

By measuring the moment of force required to twist a given wire through a measured angle, the coefficient of rigidity of the substance of which the wire is made may be determined by the use of this formula.

RIGIDITY

Steel.......	82×10^{10} dynes per sq. cm.,	12.0×10^6 lbs. per sq. inch.
Brass.......	38×10^{10} " " " "	5.5×10^6 " " " "
Glass.......	24×10^{10} " " " "	3.5×10^6 " " " "

251. Young's Modulus. When a rod or wire is stretched by a weight the elongation is very nearly proportional to the stretching force, and exactly proportional to the length of the wire. In this case the stress is the force per unit cross section and the corresponding strain is the elongation per unit length. The elasticity of stretch for the substance of which the wire is made, or *Young's modulus*, as it is commonly called, may be represented by Y and is determined by the ratio

$$Y = \frac{\frac{F}{S} \quad \text{stress}}{\frac{e}{l} \quad \text{strain}} = \frac{Fl}{Se},$$

where F represents the stretching force, e the elongation of the wire, l its length, and S its cross section.

YOUNG'S MODULUS

Copper.......	12×10^{11} dynes per sq. cm.,	17×10^6 lbs. per sq. inch.
Brass.........	11×10^{11} " " " "	16×10^6 " " " "
Iron..........	19×10^{11} " " " "	27×10^6 " " " "
Steel.........	22×10^{11} " " " "	32×10^6 " " " "
Tungsten......	41×10^{11} " " " "	59×10^6 " " " "

252. Beams. When a floor beam sags under a load the upper part is compressed and the lower part stretched. But the resistance to longitudinal stretching or compression is measured by Young's modulus, so that the stiffness of a beam is proportional to the modulus of elasticity of the material of which it is made.

In case of a beam supported at both ends and loaded at the middle, the sag or deflection y at the middle is expressed by the formula

$$y = \frac{Fl^3}{4h^3bY};$$

where l is the length of the beam, b is its breadth, and h its height, F is the load, and Y is Young's modulus for the material of the beam.

From this it appears that a beam having twice the breadth of another would sag half as much, other things being equal, while if its depth were twice that of the other it would only sag one-eighth as much. For this reason floor beams are placed on edge, the breadth having but little influence on the stiffness compared with the depth.

253. Viscosity. When a solid is strained beyond its elastic limit the strain may go on increasing indefinitely at a rate which depends on the stress to which it is subjected. Most metals show a viscosity of this kind when the distorting force is great enough, as seen in wire drawing and in the making of lead pipe. A strip of lead when stretched with a moderate weight will continue slowly elongating year after year. A glass fiber fastened at one end and having a small twisting force applied at the other will twist more and more as time goes on.

But when a substance yields continuously in this way to the *very smallest forces*, as in case of tar, pitch, or syrup, it is said to be a *viscous fluid*. In such a fluid one layer slides over another with a velocity which depends on the stress and on the viscosity, the slower the motion for a given stress the more viscous the substance is said to be.

Viscosity may be considered a kind of internal friction between contiguous layers, and the energy spent in overcoming it appears as heat.

All known liquids and even gases are more or less viscous,

and in consequence energy is spent in heat and there is loss of pressure (§ 231) whenever a fluid flows through a long pipe. The outer layers next the wall of the pipe are nearly stationary in such a case, while the velocity of flow increases toward the center or axis of the stream.

The usefulness of a lubricating oil depends, among other things, upon its viscosity. If not viscous enough it will be squeezed out of the bearing, while if too viscous it will offer needless resistance to the motion.

The viscosity of a fluid may be determined from the time required for a given quantity to escape through a long tube of small diameter, or it may be found by an apparatus called a viscosimeter in which a long inner cylinder is supported by a torsion wire in the axis of an outer cylindrical tube which can be rotated. The space between the two tubes is filled with the oil or other liquid to be tested and the torsion effect on the inner cylinder is measured when the outer one is turning at a constant rate of speed.

The viscosity of a substance depends on its temperature, and it is noteworthy that *heating a liquid makes it less viscous, while the opposite is true of gases.*

254. Energy Absorbed by Viscosity. The absorption of energy through viscosity is well shown by the following experiment, due to Lord Kelvin (Sir William Thomson).

Take two eggs, one raw and one hard boiled, and suspend each like a torsion pendulum by means of a fine wire attached to a wire sling enclosing the egg, the long axes of the eggs being vertical; then give each egg a turn or two and let it go. The boiled egg will continue oscillating for a long time, while the raw egg will almost immediately come to rest. The oscillating motion of the shell is so rapid that the inner layers of the raw egg slip on the outer ones by their inertia, and the internal friction or viscosity of the egg causes the energy of vibration to be lost in heat.

This principle has been applied by Lord Kelvin to prevent the violent swinging of a mariner's compass, due to the motion of the ship. The compass box, hung on gimbals so that it can swing freely in any direction, is made with a double bottom, and the space between the two bottoms is partly filled with a viscous liquid, such as glycerin. After any disturbance the glycerin flowing between the two surfaces of the box transforms the energy of motion into heat, and the box is promptly brought to rest.

255. Internal Friction in Solids. When a vibrating tuning fork is placed in a vacuum and supported in such a way that no appreciable vibrational energy is communicated to its support it is found that the vibra-

tions die out at a rate which shows a considerable frictional resistance within the solid of which it is made. All solids, however little they are strained, exhibit such an internal friction which resists the strain. This internal friction is of a different nature from the viscosity of a liquid. In solid bodies vibrating at a given amplitude the energy absorbed in friction for every vibration cycle is the same, whether the vibration be performed rapidly or slowly.

DIFFUSION AND SOLUTION

256. Diffusion. If a strong solution of copper sulphate is introduced by a tube into the bottom of a tall vessel containing pure water, the denser blue solution will at first be sharply separated from the clear water above. By degrees the sulphate will be seen to steal upward into the water until in time it will be uniformly diffused throughout the liquid, just as a gas expands and fills a vessel in which it is set free, though diffusion in liquids is extremely slow.

Stirring a mixture of two liquids increases the surface through which diffusion takes place and so greatly quickens the process of complete mixture. When no diffusion takes place between the liquids they will not mix.

257. Interdiffusion of Gases. Gases diffuse into each other very freely, as shown by the following experiment. Two globes are connected together, the upper containing hydrogen and the lower carbonic acid gas. In spite of the density of the carbonic acid being 22 times that of hydrogen, it will diffuse upward and the hydrogen downward till finally a uniform mixture will fill both vessels. Each expands and fills the whole space as if it alone were present.

258. Solution of Solids in Liquids. When a solid is placed in a liquid a certain amount will be dissolved, after which no more will be taken up, and the liquid is said to be saturated. The per cent that can be dissolved depends not only on the substance, but on the temperature, solubility usually increasing with rise in temperature.

The volume of the solution is usually less than the combined volumes of the two constituents and the process of dissolving is often accompanied by a change in temperature.

Solution of Liquids in Liquids. Two liquids that diffuse into each other may either mix in any proportion, as in case of water

and alcohol, or one may only dissolve a limited amount of the other. Thus if water and ether are stirred together at a temperature of 10° C., the mass will separate into a lower layer of water containing 10 per cent of ether and an upper layer of ether containing 1⅛ per cent of water. At 10° C. ether will dissolve any amount of water less than 1⅛ per cent, and water will dissolve 10 per cent or less of ether. As the temperature is raised water will dissolve less ether, while ether dissolves more water.

259. Solution of Gases in Liquids. Some liquids, such as water, dissolve all gases more or less freely. When there is simple solution without chemical union the gas is absorbed most freely when the liquid is cold and is driven off when the liquid is heated. Thus when water is heated the absorbed air escapes in bubbles before boiling takes place.

The amount of gas absorbed by a given liquid is proportional to the pressure. Soda water is charged with carbon dioxide gas under pressure, and when the pressure is relieved the gas escapes in bubbles, causing effervescence.

The power of water to absorb various gases is shown in the following table, the figures giving the volume of gas at one atmosphere pressure absorbed by unit volume of water.

ABSORPTION OF GASES IN WATER

Gas	At 0° C.	At 15° C.
Oxygen...............	0.049	0.030
Hydrogen............	0.021	0.019
Nitrogen............	0.023	0.015
Carbonic acid gas....	1.79	1.00
Ammonia.............	1140.00	756.00

260. Absorption of Gases in Solids. Certain porous solids have a great power of absorbing gases. Boxwood charcoal will absorb 90 times its volume of ammonia and 35 volumes of carbonic acid gas. This absorption seems to be due to the condensation of a layer of gas on the surface of the body.

It is by the condensation of a surface film of gas over a body that the so-called Moser's breath figures are explained. If an engraved die lie for some time on a polished plate of metal or

glass, on removing the die and breathing on the plate the engraved image is seen.

Platinum in the porous state, known as spongy platinum, absorbs hydrogen gas so powerfully that if placed in an escaping jet of hydrogen the heat developed by the condensation is sufficient to ignite the jet.

There is what seems to be a true solution of gases in some solids discovered by Graham and called by him occlusion. By heating iron wire and then allowing it to cool in an atmosphere of hydrogen, it was found that it occluded 0.44 times its volume of the gas, while platinum occluded 4 times its volume of the same gas. It is probably in consequence of this that hydrogen readily diffuses through iron and platinum when they are red hot. The most remarkable substance in this respect is palladium, which absorbs 960 times its own volume of electrolytically developed hydrogen and at the same time expands about $\frac{1}{10}$ of its volume.

261. Colloids and Crystalloids. It was found by Graham that a diaphragm of parchment or bladder would allow certain substances in solution, such as salts, to freely diffuse through it, while other substances, such as albumen, starch, gum, glue, or gelatin, could not pass through or only very slowly. He was led, therefore, to divide substances into two classes, crystalloids and colloids (from the Greek word for *glue*), which could be separated in this way.

FIG. 134. Osmotic pressure

262. Osmotic Pressure. If a strong solution of sugar or glucose is placed in a vessel opening above in a tube and closed at the bottom with a membrane, such as a bladder, and if the whole is set in a vessel of pure water, diffusion of water takes place through the membrane, and the solution accordingly rises in the tube until the pressure of the solution in the vessel is sufficient to prevent any more water from entering. The increase pressure in the interior when equilibrium is reached is known as the *osmotic pressure* of the dissolved substance.

Such a membrane, however, is not suited for *measurements* of osmotic pressure, for sugar diffuses through it to some extent, though not as rapidly

as water, also it lacks the strength and rigidity needed to support the pressure developed.

Pfeffer showed how to form by chemical means within the substance of a porous earthenware cup or cell, a membrane which was entirely impervious to sugar in solution while it freely transmitted water. A semipermeable membrane formed in this way has great mechanical strength, and H. N. Morse who has greatly improved the Pfeffer cell has been able by its means to measure osmotic pressures as high as 28 atmospheres, or 400 lbs. per square inch.

Investigation shows that osmotic pressure is proportional to the number of molecules of the dissolved substance in 1 gram of solvent, and when aqueous solutions of different substances have equal osmotic pressures, the number of molecules of dissolved substance per gram of pure solvent is the same in one as in the other, just as equal volumes of different gases at the same temperature and pressure contain equal numbers of molecules (Avogadro's law).

Capillarity and Surface Tension

263. Capillarity. Under this head are grouped a number of phenomena depending on the force of cohesion at liquid surfaces. Some of these are the upward curvature of the surface of water where it meets the side of a glass vessel, the clinging together of light bodies floating on the surface of water, the forms of drops and bubbles, and the rise of liquids in fine hair like or capillary tubes. (Latin *capillus*, a hair.)

To understand these phenomena we must first examine how the conditions at the surface of a liquid differ from those in its interior.

264. Surface Tension. If a mass of olive oil is floated in a mixture of alcohol and water of the same density as the oil, it will gather itself up into a spherical ball. Draw it out by means of a glass rod into any long or irregular shape, and as soon as it is left to itself it returns to its spherical form, exactly as if covered with an elastic skin. When a camel's-hair brush is immersed in water the bristles stand apart as freely as in air, but when it is withdrawn they cling together by the contraction of the surrounding water surface. Wet threads cling together when drawn out of water. So also a drop of mercury resting on a table is drawn up into a smooth rounded mass by the contraction of the surface, in spite of the weight of the mercury which tends to flatten it out.

265. Cause of Surface Tension. That the particles of a liquid attract each other and are held together by a strong force of cohesion may be shown by the *water hammer*, which is a bent tube (Fig. 135) partly filled with water from which the air has all been boiled out, the tube having been sealed up while boiling so that it contains simply water and its vapor with scarcely any air. The absence of air causes the water to strike the end of the tube with a sharp metallic click when it is shaken, hence its name. If the water is all run into one arm of the tube, completely filling it. and is jarred into good contact with the sides of the tube, it may then be held in the position shown in the figure and the water will remain in the full arm in spite of the fact that the pressure of the vapor in the other branch is quite insufficient to sustain the column of water; for if a slight jar is given to

FIG. 135.   Water hammer

the tube, the liquid sinks to the same level in each branch. In this case the column of water is sustained by clinging to the walls of the tube and by the cohesion of one particle to another, and it is under *negative pressure* or *tension* as much as is a rope supporting a weight.
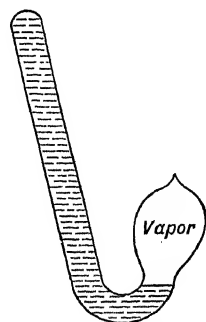
But this force of cohesion can be detected only between particles that are exceedingly close together.

The small distance within which all those particles lie that have any sensible attraction for a given particle may be called

FIG. 136

the radius of the sphere of molecular attraction. If the spheres of action are represented by the circles about A and B, it is clear that the particle at A is in equilibrium, so far as the cohesive forces are concerned, being equally drawn in all directions, while B, which is nearer to the surface than the radius of the sphere of action, has the downward attraction of liquid below the line ab balanced only by the upward attraction of the medium above the liquid surface. If the latter is a free surface with air or vapor above, the downward attraction will be in excess and the tend-

ency is to drag particles away from the surface and into the interior of the liquid. In consequence of this the surface tends to contract. If the upper surface of the liquid were in contact with a substance whose attraction for the particle B was greater than that of the liquid below ab, the upward attraction would be in excess and the surface would tend to enlarge. Thus a drop of oil on a clean glass plate will spread out over the whole surface of the glass.

Some idea of the size of the spheres of attraction may be formed from the study of soap films.

Let figure 137 represent a section of a soap film in which the circles indicate the spheres of attraction. Particles in the middle

FIG. 137

of the film, between the two dotted lines, are farther from the surface than the radius of the sphere of action, and no change in the contractile force of the film is to be expected so long as the two surfaces are thus comparatively independent of each other, but if the film is made thinner than the diameter of the sphere of action a change in the tension of the film is to be expected. Such a change — a sudden decrease in tension — is observed when the thickness of a soap film is about 100 millionths of a millimeter, indicating that the radius of the sphere of action in the soap solution is about 50 millionths of a millimeter.
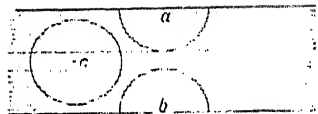
266. Measure of Surface Tension. *The surface tension of a liquid is the force with which the surface on one side of a line one centimeter long pulls against that on the other side of the line.* Thus it is the

FIG. 138

contractile force which a square centimeter of surface (Fig. 138) exerts on each of its bounding sides. In the diagram the arrows marked T indicate the outward forces due to the tension of the surrounding surface required to balance the contractile force of the surface inside the square.

267. Surface Energy. If a strip of surface one centimeter wide is made one centimeter longer than at first, the work done is measured by the contractile force or surface tension T multiplied by the distance that it is drawn out. Thus if it is drawn

out one centimeter, increasing the area of the surface by one square centimeter, the work expended is T ergs. This work is stored up as energy of the surface and is expended when the surface contracts. *Particles of liquid near the surface have thus more energy than particles in the interior, and the increase in energy in ergs per square centimeter of surface is numerically the same as the surface tension in dynes per linear centimeter.*

The following table gives some values of surface tensions at 20° C.

SURFACE TENSIONS IN DYNES PER CENTIMETER

SUBSTANCE	AIR	WATER	MERCURY
Water.........	73.5	.....	412
Mercury.......	539.0	412.0	...
Olive oil.......	34.3	20.6	335
Alcohol........	24.5	.....	...
Ether..........	17.6	.....	...

The tensions just given are for the interface between the substance at the top of the column and the one at the side.

268. Variations of Surface Tension. The surface tension of a liquid depends on temperature, in general being less with higher temperatures; it is also in case of water greatly affected by impurities. Pour a little water into a flat-bottomed porcelain tray, but not quite enough to cover the bottom. If a few drops of alcohol are added at any point, the water will rush away from that spot in every direction leaving the porcelain surface bare. If a drop of ether or alcohol is held near the surface of clean water on which lies some lycopodium powder, this will be dragged away from near the drop. The surface tension of the liquid is weakened by alcohol and even by the vapor of alcohol or ether and the surrounding uncontaminated liquid with its greater surface tension contracts and draws the other after it. In the same way are to be explained the lively movements that are noticed when minute fragments of camphor are dropped on clean water. The surface tension is weakened by the impurity at the first point of contact and as the liquid is drawn away from that point the camphor also moves, leaving a trail of contaminated surface behind it.

269. Pressure Due to Surface Tension. The contractile force of a curved surface produces a pressure inward on the concave side. Let the figure represent a drop of water which we may imagine free from gravity. It will at once assume a truly spherical form and the liquid will be under pressure in consequence of the tension of the surface.
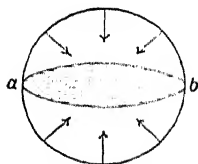
FIG. 139

To determine the amount of this pressure consider the drop as in two halves separated by the plane *ab*. All around the circumference of this section the surface tension is pulling the two halves together. The amount of this force will be $2\pi r T$ if r is the radius of the drop and T the tension.

But this force is balanced by the pressure of one half of the drop against the other. Calling p the pressure per square centimeter, the total pressure is $\pi r^2 p$. But since these two forces balance we have

$$2\pi r T = \pi r^2 p;$$

and therefore

$$p = \frac{2T}{r}.$$

Notice that this pressure increases as the size of the drop diminishes, and in a water drop one-hundredth of a millimeter in diameter it amounts to 150 grams per square centimeter.

This same expression gives the pressure, due to surface tension, of the air or vapor inside of a bubble in a mass of liquid.

In a soap bubble there is a thin film of the soap solution having two surfaces to be considered. Each of these has contractile force and consequently the total pressure inside of the bubble due to surface tension is

$$p = \frac{2T}{r_1} + \frac{2T}{r_2},$$

where r_1 and r_2 are the radii of the outer and inner surfaces of the bubble, respectively. Since these radii are practically equal, we have

$$p = \frac{4T}{r}.$$

270. Contact Angle. When a clean plate of glass is dipped into water the liquid rises in a curve against the glass. The free surface of the water is here enlarged in spite of its contractile force by the *expanding force* of the surface of contact between the water and glass.

This expanding force or *negative* tension is due to the great attraction between the water and glass, as explained in § 265. Let E represent the amount of this expansive force or negative tension of the surface between water and glass, and let T represent the tension of the water surface. Then it is clear that the liquid will rise until the upward and downward forces are in equilibrium; that is, until $E = T \cos x$.
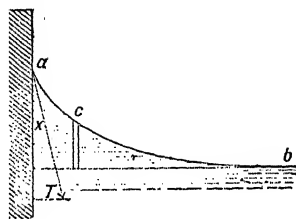
FIG. 140. Contact angle

The particular angle x at which equilibrium takes place is known as the contact angle for the two substances involved. In case of kerosene oil and glass E seems to be greater than T, and hence there can be no equilibrium, the angle x becomes 0°, and still the edge of the oil creeps up. It is in this way that a film of oil spreads over the whole surface of a glass lamp.
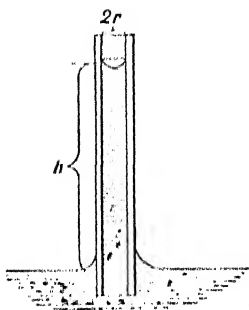
FIG. 141. Rise in capillary tube

The case of mercury and glass is different, the mercury curves *down* instead of *up* at the line of contact and the angle x is about 140°, indicating that the surface between mercury and glass has a positive tension, or contractile force.

271. Rise in Capillary Tubes. If a large glass tube, say, 2 in. in diameter, is introduced into water, the water rises around the edge on the inside until the weight of the water raised above the original level is just equal to the total upward pull of the contracting surface. This upward pull is equal to the tension per unit length multiplied by the inner circumference of the tube or $2\pi rT$, where r is the radius of the tube and T the surface tension.

If small tubes are taken the weight of liquid required to balance

this force cannot be secured without raising even the middle of the liquid above the original level, and the smaller the tube the greater the height to which the liquid will rise. Let h represent the average height of the column, then $r^2\pi h$ the volume of liquid raised; and if d is its density, its weight in grams will be $r^2\pi h d$ and its weight in dynes $r^2\pi h d g$, and this is equal to the sustaining force $2\pi r T$.

$$2\pi r T = r^2\pi h d g,$$

$$h = \frac{2T}{rdg}.$$

Hence *the height to which liquid rises in a capillary tube varies inversely as the radius of the tube*, a relation known as Jurin's law.

Of course if the contact angle is not zero, we must write $T \cos x$ in the above formula instead of T.

The pressure within the liquid in a capillary tube is less than the atmospheric pressure at all points above the level surface of the liquid in the open vessel, decreasing according to the hydrostatic law, toward the top. The curved surface at the top exerts a back pressure against the atmosphere equal to the pressure of a column of liquid of the height h.

The height at which a liquid will stand in a capillary tube is independent of its shape, depending only on the size of the tube at the point where the curved surface or meniscus stands.

A liquid cannot rise and overflow the top of a capillary tube, however short it may be, for as it reaches the top the curvature of the surface changes, becoming less until its upward pressure is just balanced by the column of liquid below.

The rise of oil in lamp wicks and of sap in vegetable fibers are familiar instances of capillary action.

272. Depression of Liquids in Tubes. In cases where the contact angle is greater than 90°, as between mercury and glass, the surface of the liquid is rounded upward in a small tube and the level of the liquid is depressed. In this case the surface being concave downward produces a downward pressure. Thus in a barometer the mercury column stands lower than it would normally do unless the tube is so large that the center of the column is sensibly flat. (See § 199.)

273. Effect of Curvature of Surface. In a conical tube a drop of water will be concave outward at both ends, but since the smaller surface has the smaller radius of curvature, it will exert the greater pressure against the air and the drop will move toward the small end of the tube, while a drop of mercury which rounds outward at both ends will be driven toward the larger end (Fig. 142). Consequently a drop of mercury will be in stable equilibrium at a widening in a narrow tube, while a drop of water will seek the narrowest point (Fig. 143).

If a capillary tube is connected at the bottom with a larger vessel of water (Fig. 144), when the water in the large vessel is at *A*, level with the top of the capillary tube, the surface of the

DROP OF MERCURY IN TUBE

DROP OF WATER IN TUBE
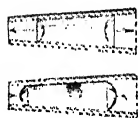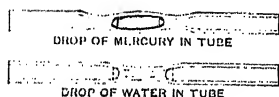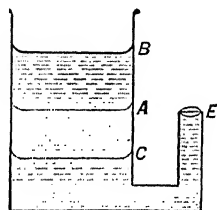
FIG. 142   FIG. 143   FIG. 144   FIG. 145

water at *E* is flat. If the level is raised to *B* in the large vessel, the surface at the end of the capillary tube will round up, taking exactly the curvature necessary to balance the hydrostatic pressure due to the height of *B* above *E*. On the other hand, if the level is lowered to *C* the surface at *E* will become concave with a curvature that will give an upward pressure equal to that of a column of water from *A* to *C*.

When a narrow glass tube is dipped in water and withdrawn a short column of water will be held in the lower end, the lower convex surface acting together with the upper concave one to support the liquid.

274. Small Floating Bodies. When a floating body is wet by a liquid, the liquid rises around it and drags it down by the weight of this raised mass. So a hydrometer with a glass stem around which the liquid rises will sink lower than it otherwise would.

The liquid curves down toward bodies which are not wetted so that the body is buoyed up by the weight of the liquid displaced in consequence of the curvature. For instance, if a clean needle is laid carefully on water it will float. The liquid is bent down where it meets the needle and therefore the volume of water displaced is much greater than the volume of the needle itself and so the weight of the displaced water may be equal to the weight of the needle.

Fig. 146

275. Attraction and Repulsion. When the liquid rises around floating bodies they are drawn together as soon as they are near enough for the curvature of the surface due to one to affect the other. Notice how small floating pieces of wood or cork cling together as soon as they are wet, so also bubbles in a cup of cocoa cling together, and are also drawn to the sides of the cup where the surface curves up.

If the cup is filled to the brim and enough added to make the surface curve down slightly at the edges, the bubbles will at once rush away from the edge.

Bodies which are not wetted and around which the surface curves down are also drawn together, but are driven away from bodies around which the surface curves up.

276. Explanation of Attraction of Floating Bodies. When two small floating objects are both wetted the liquid rises higher between them than it does on the outside, as shown in the upper part of figure 147, and since the pressure at any point in the liquid higher than the level surface is less than the atmospheric pressure, the pressure on the outside is greater and they are forced together.
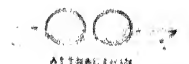
Fig. 147

In the second case the liquid stands higher around the floating bodies on the outside than it does between them, and since the pressure in the liquid at points below the level surface is greater than the atmospheric pressure, they are pushed toward each other in this case also.

But when the liquid wets one and not the other, the surface is lowered on the inside of the wetted one and raised on the

inside of the other so that in each case the pressure on the inside is greater than on the outside and the two are urged apart.

277. Soap Films. Some most interesting illustrations of surface tension are found in the phenomena of soap films. When a loop of thread is laid on a soap film formed in a wire ring and the film is broken inside the loop, the latter will be drawn into an exact circle, for it is pulled equally in every direction by the contracting film. And this circular loop may be moved from one part of the film to another without changing shape, showing that the tension does not depend on the width of the film.

If wire frames forming the outlines of cube, tetrahedron, or
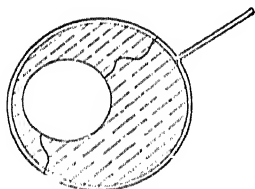
FIG. 148. Loop in soap film   FIG. 149. Cylindrical bubble
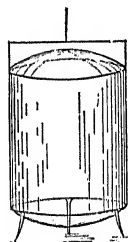
cylinder are dipped into a soap solution and then carefully withdrawn, symmetrical figures of great beauty are formed by the films.

By blowing a bubble between two rings and then drawing the rings apart until it becomes cylindrical, the ends will be seen to bulge out, showing that the air within is under pressure, and the radius of curvature of the spherical ends will be found to be twice that of the cylindrical surface. Why is this?

278. Equilibrium of Cylindrical Film and Formation of Drops. If such a cylindrical film is short relative to its length, it is in stable equilibrium; but if it is longer than its circumference, it is unstable and will collapse at one end and bulge out at the other because by so doing the surface will become smaller. This will result in its breaking into two bubbles, a large and a small one (Fig. 150), with a very small one between them; a result which is of interest, because it illustrates why a thin jet of water breaks up into drops. Imagine a thin cylindrical jet escaping from a

vessel of water. It is under pressure sidewise from the contractile force of the surface, but since it is long enough to be unstable it yields, becoming first undulatory, as shown in the upper part of the figure, then finally breaking up into alternative big and
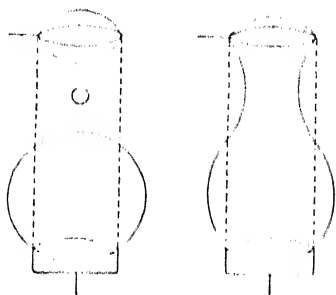
FIG. 150.   Breaking of unstable cylindrical bubble

FIG. 151.   Jet breaking into drops

little drops which are elongated when first separated and vibrate from this to the flattened form, finally settling down to spherical shape. These details were first made out by Savart; they may best be studied from photographs made when the stream is instantaneously illuminated by an electric spark.

PROBLEMS

1. Find the diameter of a drop of water in which the pressure is twice the atmospheric pressure on its surface, taking surface tension of water as 74 dynes per cm.

2. Two flat glass plates, 10×10 cm., placed face to face in a vertical position and separated only by bits of tinfoil, have the lower edges immersed in water which rises and fills the space between the plates. Find how much less the average pressure is between the plates than on the outside and thence find the force with which they are pressed together.

3. In case of a soap bubble 5 cm. in diameter, how much greater is the pressure within the bubble than without? Take surface tension 70 dynes per cm. Give answer in dynes and also in grams per sq. cm.

4. A glass hydrometer having a stem 8 mm. in diameter floats in water. With what force due to surface tension of water wetting its stem is it pulled downward?

5. How much deeper will the hydrometer in the last problem sink than if it had floated in a liquid of the same density that did not rise on its stem?

Ans. 3.8 mm.

KINETIC THEORY

279. Kinetic Theory of Gases. It was shown by Daniel Bernoulli (1700–1782) that the pressure of a gas could be best explained as due to the impacts of its molecules against each other and the walls of the vessel. In recent years Clausius and Maxwell especially have developed this theory, showing that the characteristic properties of gases are in harmony with it, and it is now generally accepted as giving a true conception of their structure.

In this theory it is assumed that the molecules of gas are constantly striking against each other or the walls of the vessel and rebounding. When two molecules approach each other, at a certain distance they experience a repulsive force which increases as they come nearer until further approach is stopped by the force and they are repelled apart or rebound. The distance between their centers when they are nearest together and about to rebound is called the diameter of the molecule. The molecule on rebounding soon gets out of the influence of the other and then flies in a straight line until it meets another from which it rebounds, either directly or glancing off sidewise, changing both its own motion and that of the molecule against which it strikes, and so it continues its path zig-zagging about. The average distance that a molecule travels between two successive impacts is called its *mean free path*. The velocity of a particular molecule is doubtless changed at every impact not only in direction, but in amount, sometimes increased and sometimes diminished, but there is no loss of energy on the whole; whatever one molecule loses the one impacting against it gains. The average kinetic energy of the molecule, and consequently its average velocity, remains unchanged unless energy is in some way communicated to the gas from outside.

280. Pressure of a Gas. An expression for the pressure of a gas may be deduced in an elementary way by neglecting the *size* of the molecules and their impacts against each other and considering each molecule as rebounding only from the walls of the vessel. Imagine a cubical vessel one centimeter each way, and for simplicity conceive the whole number of molecules N contained in it to be divided into three equal groups, one group

rebounding between the sides AD and BC and producing pressure against them, the other groups being directed against the other pairs of sides. If V is the velocity of a molecule, it will strike against the side BC once every time that it travels across the vessel and back again, a distance of 2 cms. The number of impacts per second of one molecule against the side BC will therefore be $\dfrac{V}{2}$.
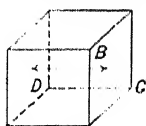
Fig. 152

The momentum of the molecule before impact is mV toward the side, after impact it is mV in the opposite direction; the total change in momentum of the molecule in one impact is $2mV$ where m is the mass of the molecule. But there are $\dfrac{V}{2}$ impacts per second, so that each molecule in rebounding from one side experiences a change of momentum per second $2mV \times \dfrac{V}{2} = mV^2$ and since the whole number of molecules impacting against the side BC is N the total *change in momentum produced by that side in one second is* $\frac{1}{3}NmV^2$, and this is, therefore, the *force* against the side. But since the side BC has unit area, the force against it equals the pressure, hence

$$p = \tfrac{1}{3}NmV^2 \qquad (1)$$

where p represents the pressure of the gas.

Now, the product Nm is the total mass of gas in one cubic centimeter, or its density d, and hence:

$$\frac{p}{d} = \tfrac{1}{3}V^2. \qquad (2)$$

Maxwell's Law. It has been shown on mechanical grounds, by Maxwell and others, that *when two masses of gas are at the same temperature, the average kinetic energy of a molecule of the one is equal to the average kinetic energy of a molecule of the other.*

That is,

$$\tfrac{1}{2}m_1V_1^2 = \tfrac{1}{2}m_2$$

where m_1 and m_2 are the masses of the molecules of the two gases and V_1 and V_2 are their velocities.

Boyle's Law. According to the law just stated the kinetic energy of the molecules in a mass of gas is determined by its temperature, and hence V changes only when the temperature of the gas changes. *Formula (2) above, then, is in agreement with Boyle's law and expresses the fact that the pressure of a gas is proportional to its density when the temperature is constant.*

This formula may be used to calculate the average molecular velocities, giving as follows:

MEAN VELOCITY OF MOLECULES IN GASES AT 0° C.

Hydrogen...................... 1843 meters per sec.
Nitrogen .... 492 " " "
Oxygen......................... 461 " " "
Carbon dioxide................. 392 " " "

Avogadro's Law. When two different gases have *the same pressure* we have by equation (1)

$$\tfrac{1}{3}N_1 m_1 V_1{}^2 = \tfrac{1}{3}N_2 m_2 V_2{}^2 \tag{4}$$

If the two masses of gas are also at *the same temperature,* we have by (3)

$$\tfrac{1}{2}m_1 V_1{}^2 = \tfrac{1}{2}m_2 V_2{}^2 \tag{5}$$

and combining the two equations we find

$$N_1 = N_2;$$

that is, *the number of molecules per cubic centimeter is the same in all gases at the same temperature and pressure.* This is known as Avogadro's law, and was reached by him from purely chemical considerations.

281. Molecular Magnitudes. The number of molecules in a gas at one atmosphere pressure and 0° C. is found to be 27.05×10^{18} per cubic centimeter, or 443 million million million per cubic inch. Several different methods lead to approximately this result, but the most accurate determination is by an electrical method to be explained later (§ 628).

The *mean free path* or average distance that a molecule travels before striking against another may be deduced by the kinetic theory when the viscosity of the gas is known.

Also by several different lines of reasoning that cannot here be discussed the effective diameters of gaseous molecules have been approximately determined.

MOLECULAR MAGNITUDES IN GASES AT ATMOSPHERIC
PRESSURE AND 0°C.

	Mean Free Path	Diameter of Molecule in Millionths of a Millimeter			
Nitrogen..........	.0000098 cm.	0 28	27 05	10⁶ – 10⁵	10⁶
Hydrogen..........	.0000185 cm.	0 21			
Carbon dioxide....	.0000068 cm.	0 37			

The numbers representing the diameters must be regarded as only approximations to the truth, but which doubtless express the true *order* of magnitude, being probably neither 10 times too large nor too small.

In case of nitrogen at atmospheric pressure the mean free path is about 350 times the diameter of the molecule. More than one million such molecules in a row would be required to make a length of 1 mm. Lord Kelvin has estimated that if a drop of water were magnified to the size of the earth, " the structure of the mass would then be coarser than that of a heap of fine shot, but probably not so coarse as that of a heap of cricket balls."

282. Brownian Movement. The English botanist, Brown, in 1827, on observing with the microscope very fine particles held in suspense in a mass of water, discovered they were in constant irregular motion, and the smaller the particle the more lively was the motion observed. It is a spontaneous motion that never ceases, and is believed to be caused by the incessant motion of the molecules of the liquid, which bombard the particle on all sides driving it hither and thither.

The French physicist, Perrin, has made a careful study of this phenomenon using an emulsion in water of exceedingly fine grains of mastic. He finds by exact measurement of the distribution of the grains and the amount of their motions that they distribute themselves just as should be expected from the

kinetic theory, and even deduces by inference from his measurements the number of molecules in a cubic centimeter of gas under standard conditions, finding 30.5×10^{18}, in good agreement with determinations by other methods.

REFERENCES

C. V. Boys: *Soap Bubbles.*

J. Perrin, translated by F. Soddy. *Brownian Movement and Molecular Reality.*

WAVE MOTION AND SOUND

SURFACE WAVES

283. Wave Motion. The phenomena of wave motion are perhaps most easily grasped from the study of water waves. Looking upon a series of waves coming across a smooth lake from a passing vessel, we notice the steady advance of a definite form of motion, having a velocity which is quite independent of that of the vessel, and carrying energy, for the water in a wave is in motion and has kinetic energy.

The water over which the waves have passed is left calm, and if a floating cork is observed it will be seen to rise and move forward with the crest of the wave, then sink and move backward in the trough, repeating the motion with the next wave, and coming to rest in its original position when the disturbance has passed. The motion of the cork, which is that of the water in which it lies, shows that the wave does not carry along with it a mass of water, but that the motion and energy are passed along from one mass of liquid to the next.

Fig. 153.   Origin of a water wave

A wave may be defined as a form or configuration of motion advancing with a finite velocity through a medium.

By means of waves energy is transmitted, being passed along from one part of a medium to the next by the interaction of adjoining parts.

284. Origin of a Series of Water Waves. When a stone is dropped into a smooth pond, water is carried down by the motion of the stone, as shown at A, figure 153, and also thrust out in a ridge at B; beyond C the surface is undisturbed. Then it begins to rush back toward A from the surrounding parts and B sinks, at the same time the forward part of the wave between B and C is urged forward, in part by its forward momentum and in part by the pressure. The rush toward A does not stop

when A has risen to the level, but continues until the kinetic energy of the flow toward A is spent in heaping up water at A at the expense of the hollow at B, as shown in 2, figure 153. Thus there are set up oscillations at A, the energy of which is gradually spent in sending out waves. *Each wave takes with it a certain definite quantity of energy which remains with it as it advances*.

285. Motion in a Water Wave. In a series of water waves the motion of the particles is as shown in figure 154, each particle moving around in a closed curve, which in the simplest form of wave is a circle. In the diagram are shown the paths of motion of nine water particles which were originally equidistant and one-eighth of the whole wave length apart. Each moves clockwise in a circle and all with the same uniform velocity; but while particle a is at the top of its path, b is back of the top by one-
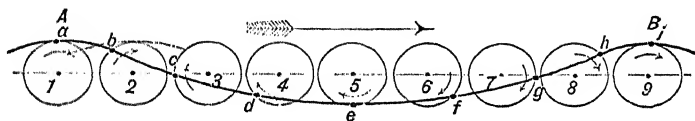
FIG. 154. Motion in a simple oscillatory water wave

eighth of a circumference. The position of each in its path is called the *phase* of its motion; a and i are said to be in the *same phase*, while the phase of e is opposite to that of a and i; also c and g are opposite in phase because they are a half-circumference apart in their motion.

Each particle in the diagram differs in phase from the next one by one-eighth of a complete revolution. Of course all the particles which were between a and b in the undisturbed condition of the surface will still be between a and b and will have intermediate phases, thus forming the surface of the wave between those points.

It will be seen that the wave is advancing in the direction of the long arrow at the top, for an eighth of a period later b will be at the top and a will have passed beyond, and the position of the crest of the wave will be as shown by the dotted line. In the time of a complete revolution or *period* of the particles the wave will advance from a to i and a new crest will have come to a.

The *wave length* is the distance between particles in the same phase of motion, in this case from *a* to *i*.

The *amplitude* of the wave is the radius of the circles, which is the distance that a particle is displaced from its equilibrium position; it is one-half the vertical height of the wave from trough to crest.

The *velocity* of a wave is the velocity with which a particular phase of motion moves along; for example, it is the velocity with which the crest of the wave moves along. Since a wave travels the whole wave length λ, in the period of revolution of a particle T, we have

$$V = \frac{\lambda}{T}$$

where V represents the velocity of the wave.

The *frequency* of a series of waves is the number passing a particular point per second. If n waves pass per second there must be n complete waves in the distance V traversed by the waves in one second, or

$$V = n\lambda.$$

The velocity of the particle in its circular orbit must not be confused with the velocity of the wave. It is always less than the wave velocity and depends on the amplitude of the wave. A mechanical wave model devised by Lyman exhibits the motions in figure 154 and admirably illustrates the motion in water waves.

286. Decrease of Amplitude with Depth. In consequence of their difference of phase, the particles *a* and *b*, near the crest of

FIG. 155.   Amplitude of wave decreases with increased depth

the wave, are nearer together than their positions of equilibrium, while near the trough of the wave *e* and *f* are farther apart. But since water is practically incompressible its volume does not change and a little mass of water which was originally of cubical form must become elongated vertically near the crest and horizontally near the trough as shown in figure 155. The water par-

ticles on the lower surface of the cube must, therefore, have less amplitude of motion than those at the surface, as is evident from the figure. At a depth equal to the wave length the amplitude is only $\frac{1}{535}$ of that at the surface.

287. Velocity of Oscillatory Wave. The type of water wave described is known as a simple oscillatory wave and the varieties of form observed in ocean waves are due to a number of such waves superposed. In a smooth pond it may easily be observed that two independent series of waves may cross each other, producing a complicated resultant motion; but after they have passed, each is found to have kept its original motion undisturbed.

It is shown in treatises on hydrodynamics that the velocity of a simple oscillatory surface wave is expressed by the formula

$$V = \sqrt{\frac{g\lambda}{2\pi}}.$$

From which the following velocities are found:

Wave length	Velocity
75 ft.	13.3 miles per hour.
300 ft.	26.6 miles per hour.
1200 ft.	53.2 miles per hour.

From this it is clear that a group of waves, as we see them in the ocean, resulting from the superposition of a variety of waves of different lengths, must continually change in form, as the component simple waves travel with different velocities.

288. Ripples. The formula just given for the velocity of a wave assumes that forces due to the weight of the liquid are the only ones involved. But the surface tension of the liquid also plays a part, though it is entirely insignificant except in very short waves or *ripples*. The complete formula for the velocity of a wave is

$$V = \sqrt{\frac{g\lambda}{2\pi} + \frac{2\pi T}{\lambda d}}$$

where T represents the surface tension of the liquid and d its density. The effect of surface tension is to increase the velocity of shorter waves. When water waves are about 17 mm. long

their velocity is a minimum, being 23.3 cms. per second.   Longer waves travel faster because gravitation force predominates, while in shorter waves surface tension has the principal effect.

COMPRESSIONAL WAVES

289. Compressional Waves.   Water waves of the type just considered are *surface* waves, and can only exist at the surface of a medium.   But the kind of wave now to be studied can travel in every direction through an elastic medium.

Consider the model shown in figure 156, which represents a series of equal masses resting in a frictionless groove and con-

FIG. 156.   Illustrating elasticity and inertia of medium

nected by springs.   If the first mass is moved toward the second, the latter will move *because* the spring between the two is compressed. *But it will not begin to move until after the first mass has approached it;* for *if* the two moved exactly together there would be no compression of the spring between them, and consequently no force exerted on the second mass to move it.   As the second mass moves forward there is compression of the second spring, followed by motion of the third mass, and so on, the masses being set in motion one after the other as the wave of compression reaches spring after spring.

So also if the first mass is drawn away from the others, the first spring is stretched, causing motion of the second mass which stretches the second spring.   The motion is therefore communicated through the whole series as a wave accompanied by stretching or expansion of the springs.

Such a wave of expansion or compression is set up whenever a material object is set in motion or brought to rest; for all bodies may be considered as made up of massive particles in elastic equilibrium with each other, like the balls and springs in the diagram.   Thus, when a chair is lifted, a wave of expansion runs down through it, and when it is set on the floor a wave of compression runs up.

290. Newton's Formula for Velocity of a Compressional Wave. The velocity of such a wave depends on the elasticity and density of the medium.   Recurring to the illustration, it is evident that making the springs between the balls stiffer will increase the

speed with which the motion will be communicated from one ball to the next, while if the masses of the balls are made greater the effect will be to make the speed of the wave less.

It was proved by Newton from the principles of mechanics that *the velocity of a wave of compression or expansion in a medium of which the volume elasticity is E and the density d is expressed by the formula,*

$$V = \sqrt{\frac{E}{d}}.$$

A simple proof of this formula will be found in Appendix **II** at end of book.

291. Motion in a Series of Compressional Waves. If the first of the series of balls represented in figure 156 is made to oscillate regularly backward and forward, now moving toward the second ball and now away from it, a *series* of waves will be sent along the row of balls, alternately waves of compression and expansion; and each ball will oscillate just as the first one does, though the second will always be in a phase of motion a little behind the first, the third will lag behind the second and so on.
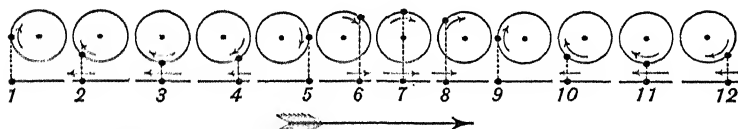
FIG. 157.   Motion of particles in a sound wave

This is precisely the kind of motion which is set up in air by a tuning-fork or other rapidly vibrating body, and which excites in our ears the sensation of sound. Such waves in air are accordingly known as *sound waves*.

The details of the motion in a sound wave may be understood from figure 157.

The diagram illustrates the relative phases of motion of a series of particles in the wave one-eighth of a wave length apart. Each particle is shown as a black dot on a short straight line which represents the path in which it oscillates, the center of the line being its equilibrium position. Suppose the first particle is made to oscillate in simple harmonic motion, then that will be the mode of vibration of all the particles, and each will move back and for-

ward keeping vertically under an imaginary companion particle that is supposed to move with uniform velocity around in the corresponding circle shown in the diagram.

It will be seen from the positions of the companion particles in the circles that, taken in the order in which they are numbered, each is one-eighth of a complete vibration behind the phase of the preceding particle; indeed, the associated circles are only used to show the relative phases of the numbered particles below, which represent actual particles in the medium.

Particle 1 is at the end of its path, while the second particle is moving toward the end and will be there an eighth of a period later, when 3 will be in the phase now shown by 2, and so on; therefore the wave will have moved forward in the direction of the long arrow underneath. It will be seen that particles 1 and 9 are in the same phase, and accordingly the distance between them is the *wave length*. The particle at 7 is in the center of a condensed region, where the particles are closer together than normal, while those at 3 and 11 are the centers of rarefied or expanded regions. *In the condensed region the particles are moving forward in the direction in which the wave is advancing; in the rarefied region they are moving opposite to the wave.* There are intermediate points where
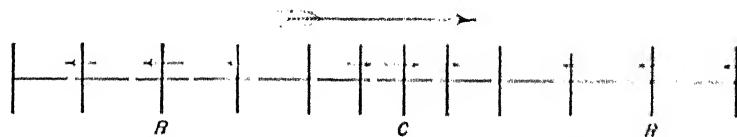
Fig. 158.   Motion of air layers in sound wave

the medium is neither condensed nor rarefied, where the particles are for the instant at rest at the end of their paths of vibration, as at 1, 5, and 9.

It must not be forgotten that each of the particles considered is only one of a *layer* all vibrating in the same way. This is represented in figure 158, where the dots of the previous diagram are replaced by heavy lines which represent successive layers of particles, differing in phase by one-eighth of a period. The small arrows indicate the velocities of the layers at the given instant, and the instantaneous position of the regions of condensation and rarefaction are marked by the letters *C* and *R*, respectively.

To sum up, *in a compressional wave the particles vibrate longitudinally, or back and forth in the direction of advance of the wave, and there is a progressive change in phase, in consequence of which alternate regions of compression and rarefaction are produced.*

The *amplitude* in such a wave is the distance that a particle oscillates on each side of its equilibrium position, or half the whole distance through which it vibrates.

292. Illustration. The propagation of a wave of compression or rarefaction may be very well shown in a regular spring a meter and a half long which is supported by threads in a horizontal position, as shown in figure 159. The turns of the spring should
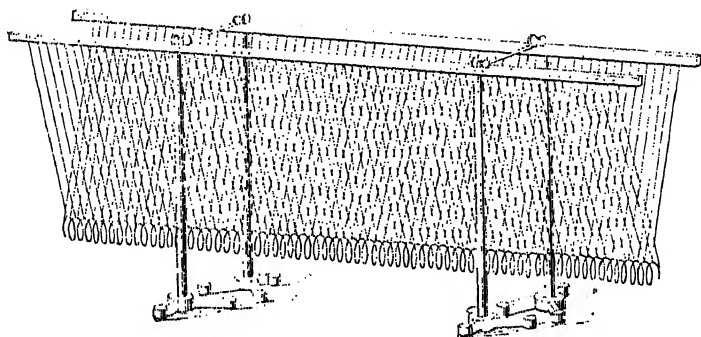
FIG. 159. Spring wave model

be rather large and it should be of such a stiffness that a wave will take a second or two to travel its length.

SOUND

293. Sound Communicated by Waves. There are three principal evidences that sound is communicated by compressional waves through material bodies. First, sound is not communicated through a vacuum; second, the motions of sounding bodies are such as might be expected to set up compressional waves; and, third, the observed velocity of sound is the same as that of compressional waves, both in air and in other media.

294. Sound Requires a Material Medium. Place an alarm bell rung by clockwork under the receiver of an air pump, as in

figure 160, so that it may rest on a mass of soft cotton, or is otherwise supported so that no vibrations can be transmitted through its supports to the plate of the air pump. When the air is exhausted from the receiver the bell is no longer heard, however vigorously it may be ringing. *Sound waves, therefore, do not pass through a vacuum, they require a material medium.*

295. Sound Originates in Vibrating Bodies. All sources of sound are vibrating bodies capable of setting up air vibrations. A brass plate supported at the center and covered with sand if set in vibration by a bow may be made to give out a variety of different sounds, but in each case there is a characteristic arrangement of the sand showing that a particular mode of vibration of the plate corresponds to each sound. (See Fig. 202.) Tuning forks
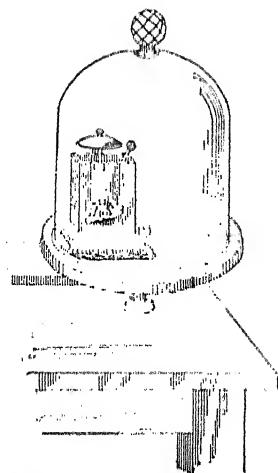
FIG. 160.   Bell in vacuo

are set in vibration by being struck, strings by being bowed, the vibrations of the string being evident to the eye or causing a buzzing sound when the string is touched with a piece of paper. In reed instruments the metal tongue of the reed vibrates strongly when sounding, and even in flutes and organ pipes it may easily be shown that the air is set in strong vibration.

296. Velocity of Sound. The velocity of sound in air was determined by two Dutch observers, Moll and Van Beck, in 1823, by timing the interval between seeing the flash of the discharge of a distant cannon and hearing its report.

Cannons were set up on two hills nearly 11 miles apart, and by observing alternately first from one hill and then from the other the observers sought to eliminate the influence of any air currents which might exist. At the same time the temperature of the air was observed at a number of points between the two stations. The velocity was thus found to be 1093 ft. or 333 meters per second at 0° C.

Regnault (1810–1887) conducted an extensive series of investigations on the velocity of sound in the Paris water mains, which

afforded large tubes free from wind disturbance and at a uniform temperature. He made use of an automatic apparatus by which the instant of discharge of a pistol was recorded electrically on the rotating drum of a chronograph, while the arrival of the sound at the distant station, where it caused a thin stretched membrane to vibrate, was electrically recorded on the same drum. By these experiments he found that the velocity was influenced by the size of the pipes to a small extent, and also that very intense sounds traveled slightly faster than feebler ones. He also conducted experiments in the open air, using the same recording apparatus, and found the velocity of sound in dry air at 0° C. to be 330.6 meters per second. More recent results give a velocity of about 331.8 meters per second in air free from carbon dioxide.

Bosscha determined the velocity of sound by causing two little hammers to give simultaneous taps at regular intervals, the frequency of the taps being determined by a pendulum which made electrical connections at every swing. If one of the sounding instruments is placed beside the observer while the other is moved away, the taps are no longer heard simultaneously, but those from the more distant one come later; if moved far enough apart so that the sound from the tap of the distant hammer reaches the ear at the same instant as the next succeeding tap of the nearer hammer, the two are again heard simultaneously, and the distance between the two sounders divided by the time interval between the taps gives the velocity of sound.

297. Velocity of Sound in Water and in Solids and Gases. Colladon and Sturm measured the velocity of sound in the water of Lake Geneva by causing a bell to sound under water and using as a receiving instrument a sort of ear trumpet with the outer end closed by a rubber diaphragm and placed beneath the surface of the lake. The velocity was found to be 1435 meters per second. According to recent results the velocity in air-free water was found to be 1441 meters per second at 13° C., and 1505 meters per second at 31° C. In solids the velocity of sound is usually measured by the longitudinal vibrations of rods or wires as explained later, § 355.

The velocity of sound in various gases and vapors has been determined by comparison with that in air by the method of Kundt, § 348.

The velocities of sound in some common media are given in the following table. The velocity of sound in wood and steel is so great that a person standing near one end of a long beam or rail that is struck at the farther end hears two sounds in quick succession, first that transmitted by the solid and then that through air.

VELOCITIES OF SOUND

Medium		Metre per sec.	
Air at 0° C. and 76 cms. pressure	{ Regnault ...	330·6	
	{ Bosscha ...	331·6	
	mean	331·1	1,086·7
Hydrogen at 0° C. and 76 cms. pressure		1,286·0	4,219·0
Carbon dioxide at 0° C. and 76 cms. pressure		261·0	856·0
Water at 13° C		1,437·0	4,715·0
Brass rod...................		3,600·0	11,800·0
Iron rod........................		4,950·0	16,240·0
Steel rod.................		5,000·0	16,410·0
Pine-wood rod (along the grain)..		3,300·0	10,830·0

298. Velocity of Compressional Waves. The velocity of a compressional wave in air may be readily calculated by Newton's formula

$$V = \sqrt{\frac{E}{d}}.$$

It was shown by Newton that the elasticity of a gas at constant temperature is equal to its pressure (see § 249). But on substituting pressure for elasticity in the above formula the calculated velocity was found to be too small.

Laplace pointed out that though the *average* temperature of air is not changed by the passage of sound waves, yet in the compressed part of a wave the air is heated for the instant, and where it is rarefied there is cooling, and that these changes take place so rapidly that there is no time for heat to flow from one part to another, so that the air is practically in an *adiabatic* condition (§ 249). The effect of heating during compression is to resist the compression, and cooling during expansion acts to oppose the expansion, the effective elasticity in this case is therefore increased and in case of air has been found to be 1.40 times as great as if the

temperature had remained constant. The formula thus becomes for a gas like air,

$$V = \sqrt{\frac{p}{d}\, 1.4}.$$

Substituting the values for air at normal temperature and pressure, and expressing both pressure and density in C. G. S. units, we have

$$V = \sqrt{\frac{76 \times 13.6 \times 980 \times 1.40}{0.001293}} = 33{,}120 \text{ cms. per sec.,}$$

which is in good agreement with the velocity of sound as found by experiment.

In case also of solids and liquids the results obtained by the formula agree with velocities obtained by direct experiment. The elasticities of these substances are so much greater than that of air that the velocities of sound in them are large in spite of their great densities.

Thus in water the elasticity or ratio of pressure increase to corresponding decrease in volume is, in C. G. S. units,

$$\frac{76 \times 13.6 \times 980}{0.000047} = 2.16 \times 10^{10} \text{ dynes per sq. cm.}$$

or 15,230 times that of air, while it has only 773 times the density of air.

299. Influence of Temperature and Pressure on Sound Velocity in Air. From the formula in the preceding paragraph it is clear that *the velocity of sound in air is independent of the pressure*, for when the pressure is increased the density increases in the same proportion, by Boyle's law, and the ratio $\frac{p}{d}$ remains constant, and consequently the velocity is constant so long as the temperature is not changed.

But if the temperature is raised, pressure being constant, the density diminishes and the ratio $\frac{p}{d}$ increases. Hence *the velocity of sound in air is increased* $\frac{1}{546}$, *or about 2 ft. or 0.60 meter per second per degree Centigrade rise in temperature.*

300. Influence of Pitch on Velocity of Sound. It may be easily noticed that the notes of music coming from a distant band are heard in the same relation to each other as if the band were near. There is no confusion of the melody such as would result if high-pitched sounds traveled faster or slower than low ones. Regnault made careful observations on this point and concluded that *the velocity of sound is the same whatever the pitch may be.*

It will be shown later that the pitch of a sound depends upon wave length, hence we conclude that *the velocity of sound is the same for all wave lengths.*

REFLECTION AND REFRACTION OF WAVES

301. Reflection of Water Waves. When a water wave meets an immovable obstacle it is turned back or reflected. Since the obstacle does not move, it cannot receive energy from the incident wave, and therefore the reflected wave carries the energy away. Each point of the obstacle reacts against the waves which meet it and so produces a periodic disturbance and may be regarded as a center from which waves are sent out. The reflected wave as a whole is the resultant of these little waves coming from each point of the obstacle.
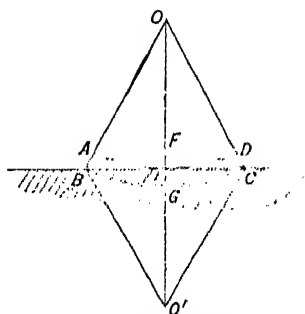
FIG. 161.   Reflection of waves

Suppose a wave from a center O, figure 161, meets a straight wall BC. When in the position AED the disturbance has just reached the wall at E and is about starting back. By the time the wave at A has advanced to B and at D has reached C, the part of the wave reflected at E will have returned an equal distance to F.

If the wall had not been there the wave would have advanced to the position of the dotted line BGC, but, since after reflection it has the same velocity as before, the reflected wave will at each point have gone back from the wall as far as it would have passed the line BC if the wall had not been there. The front of the returning wave BFC has therefore the same curvature as

BGC if the wall is *flat*. The returning wave is therefore circular having its center at a point O' which is as far back of the wall as the center O is in front of it; and the line OO' is at right angles to the wall.

Another method of looking at this subject is interesting. The effect of the vertical wall is to oppose any forward or backward motion of the water particles next to it without interfering with vertical motions. Let us now imagine the wall removed and that whenever a wave starts from O an exactly equal wave sets out from O'. The waves will meet along the line *BC* and the forward or backward movement due to the one will be exactly balanced by that of the other, while their vertical movements will be added. There results, therefore, an up-and-down oscillation along the line *BC* exactly as if the wall were there.

On each side of the line *BC* there will be waves coming toward the line and others going back from it exactly as if reflected from it. And indeed they may be properly regarded as reflected, for there is no transfer of energy across the line *BC* because there is no forward or backward motion across that line, and if a thin wall were slipped in along *BC* separating the two systems of waves the motion would not be changed on either side.

302. Angle of Reflection. When a wave front meets a reflecting surface obliquely, the direction of the wave front and its direction of propagation are changed as shown in figure 162. At W_1 is shown a portion of a wave front approaching P where it is reflected, afterward advancing as shown at W_2 as if it came from O_1. The angle i between the direction of advance of the incident wave and the normal to the surface is called the angle of incidence, while the angle r between the direction in which the reflected wave moves and the

Fig. 162

normal N is called the angle of reflection. *The angle of reflection is equal to the angle of incidence.* For the angles a and b are clearly equal and the angle i is equal to a, and r is equal to b, since the lines OO' and NP are parallel.
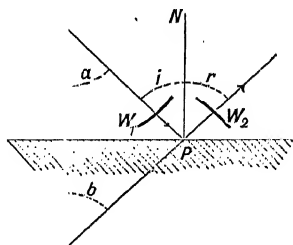
It is interesting to see how the reflected wave may be regarded as the resultant of little waves coming back from each point of the reflecting surface. Let AB (Fig. 163) be a wave front meeting the reflecting surface AC at A. The disturbance at

A causes a little circular wave to go back which will have a radius *AD* equal to *BC* by the time that the wave front at *B* has reached *C*, since the velocity of the returning wave is the same as that of the advancing one. So also the circular wavelet starting backward from *F*, when the advancing wave has reached

FIG. 163

the position *FG*, will have a radius *FK* equal to *GC* when the wave front at *B* reaches *C*.
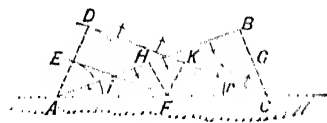
From each point in succession of the reflecting surface between *A* and *C* these elementary waves spread out just as from *A* and *F*, in arcs of circles whose centers lie on the line *AC*. The *envelope* of these circles is the line *DC* which is tangent to all of them. All these elementary waves therefore act together and combine to produce a new wave front along the line *DC*. Since *AD* is equal to *BC* and is at right angles to the tangent *DC*, just as *BC* is at right angles to *AB*, the two triangles *ADC* and *ABC* are equal and the angle of incidence *BAC* is equal to the angle of reflection *DCA*.

While the elementary waves are conceived as spreading out in circles in every direction, *they produce an effect only along the wave front DC where all act together*, for it may be shown that they *interfere* (§ 332) with each other in other directions, such as off to one side, *unless the reflecting surface is very small*. If the distance *AC* is not more than the length of a wave, regular reflection will not take place, but the reflected wave will spread out in all directions as if the reflecting surface were the center of disturbance.

303. Refraction of Waves. When waves pass from one medium into another there is generally a change in velocity,

FIG. 164

which causes the direction of the wave to change when it meets the surface of separation obliquely.

Let *AB* represent the advancing wave front, meeting at *A* the second medium where the velocity is less. While the wave front advances from *B* to *D* in the first medium, the wave from *A*

will have gone a *less* distance AC in the second medium in consequence of the smaller velocity. The new wave front will be CD, tangent to the elementary waves from A and from all points between A and D. The direction of advance of the wave is changed *toward* the perpendicular, from BD to DE. An intermediate position of the wave, where part is in one medium and part in the other, shows a sharp bend at G.

If the velocity in the second medium were *greater* than that in the first the waves would become by refraction more oblique to the surface of separation instead of less oblique as in the case illustrated.

The law of refraction will be more fully discussed in connection with the study of light.

The refraction most commonly noted in water waves is when they run obliquely into shoal water near shore, where their velocity is retarded. The effect is to swing the wave front around more nearly parallel with the shore.

304. Reflection of Sound. In the reflection of sound the same principles apply as in the case of water waves. Sound waves reflected from a large flat surface appear to come from a point as far behind the surface as the sounding body is in front of it. Echoes from buildings, cliffs, and even from a wooded hillside are familiar examples of the reflection of sound. If there are a series of cliffs or shoulders of rock at different distances multiple echoes are heard. A pistol shot from a boat on a smooth lake comes as a single sharp sound followed by faint echoes from the distant shores, but if the water is rough the shot is followed by a reverberating roar as the sound comes back reflected from wave after wave.

By means of a large parabolic mirror the tick of a watch placed at its focus is reflected so that it may be heard 50 ft. away by an observer having his ear at the point on which the reflected waves are converged. The proper position to hold the ear may be found by observing where the image is formed of a light placed at the focus of the mirror, showing that the law of reflection is the same for sound as for light.

A watch is used in this experiment because the waves of sound which it gives out are so short, even relative to the size of the mirror, that the law of regular reflection holds.

In rooms with arched ceilings focal points may sometimes be

found such that sounds going out from one point are converged toward the other. A person holding his ear at one point can hear the slightest whisper coming from the other.

305. Reflection and Refraction of Sound Waves. Whenever sound waves meet the surface between two media usually both
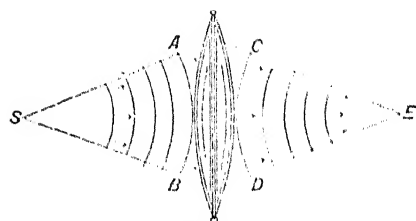
Fig. 165

reflection and refraction take place. If there is a very great change of density and elasticity most of the energy goes into the reflected wave, and the refraction will be slight. On the other hand, if the two media, like two strata of air at different temperatures, differ only slightly in their properties, most of the energy will be transmitted in the refracted waves into the second medium and but little will be reflected back from the surface.

If a lenticular bag of thin rubber is filled with carbonic acid gas (CO_2) in which sound travels more slowly than in air, the sound from the ticks of a watch will be concentrated at a focus conjugate to the position of the watch, just as light is converged by a lens of glass which retards its waves. For in passing through such a lens the middle part of the wave AB is more retarded than its edges, so that it is transformed into the form CD which is concave toward the ear at E.

Fig. 166. Sound waves changed by wind

306. Effect on Fog Signals. On account of reflection and refraction from strata of air of different temperatures, or from foggy layers, the sound of a fog horn may be entirely unheard by a vessel near the shore and in danger. If the lower portion of a horizontally moving sound wave is in warmer air than the upper part it will travel faster and cause the wave front to change its direction and may even cause it to curve upward.

So also currents of air, causing one part of a wave front to move faster than another, will change its form and consequently the direction in which it advances. Thus the observer at A (Fig. 166) might not hear the church bell when the air was still because of the screening effect of the ridge, but if a breeze were blowing which being stronger above would carry the upper part of the waves along faster than their lower part, the wave fronts would be tipped over so that they might come down to A, causing the bell to be heard.

307. Echo Depth Sounding. A useful application of the reflection of sound is in the echo depth finder. A special oscillator placed on the bottom of the hull of a ship sends an intense group of sound waves downwards which is reflected from the sea bottom and returns to the ship again, where in turn it is detected by means of a *hydrophone*. The hydrophone, which operates on the principle of the telephone (§ 738), transforms the sound into electrical oscillations which, on being amplified, cause a tube containing some neon gas to glow for a moment. The time between the starting of the wave group and its reception by the ship again is accurately measured on a dial with a pointer revolving at a known rate and which carries the small neon tube on its tip. When the pointer passes the zero mark on the dial, a contact is made which sends out the group of waves for a small fraction of a second. After the pointer has revolved a short distance beyond the zero point, a flash is observed at which position of the pointer the depth in fathoms is read directly on the dial. Since the velocity of sound in water and the velocity of the pointer are both known the scale on the dial is easily calibrated. This type of depth finder is known as the *fathometer*.

SOUND CHARACTERISTICS

308. Sound and Noise. Sounds that have a sustained and simple character and do not seem to be a mixture of various different sounds may be called *tones* or *musical sounds*. Abrupt and sudden sounds that do not last long enough to convey any idea of musical pitch, or mixtures of discordant sounds, are *noises*.

309. Tone Characteristics. A musical sound or tone has *intensity*, *pitch*, and *quality* or *timbre*, and each of these depends

upon a physical property of the sound wave. The *intensity* of a sound depends upon the amplitude or the energy of the vibration, the *pitch* depends on the frequency of the waves, and the *quality* depends on the particular manner of vibration of the particles.

310. Intensity. Intensity of a sound, as the term is ordinarily used, refers to the strength of the sensation excited by the sound wave. It depends upon the amplitude of vibration in the wave, for increasing the amplitude of vibration of the sounding body increases the loudness of the sound. But one is not simply *proportional* to the other, and if two sounds of different pitches are equally intense, it by no means follows that the amplitudes of vibration are the same. Usually the higher pitched tone will be more intense for a given amplitude than one of lower pitch.

The term intensity as applied in physics to sound waves refers to the energy of the motion and is measured by the energy transmitted per second through 1 square centimeter of surface.

Just how the *intensity of the sensation* is related to the *energy of the sound vibration* is a question for the psychologist.

The energy per cubic centimeter in a sound wave depends on the density of the medium d, the square of the frequency of vibration n^2, and the square of the amplitude a^2, and is expressed by the formula

$$E = 2\pi^2 d n^2 a^2.$$

It has been found by Lord Rayleigh that when the amplitude of vibration of the air particles is as small as one-millionth of a millimeter the sound is barely audible, while an amplitude as great as a millimeter would occur only in the very loudest sounds.

311. Decrease in Intensity with Distance. When sound waves can spread out in every direction from a sounding body forming a series of spherical waves the intensity varies inversely as the square of the distance from the source. For the same amount of energy is transmitted across every spherical surface having its center at the source of sound, and the larger the surface of such a sphere the smaller will be the energy transmitted per unit surface. The spherical wave front at C, for example, will have four times the area of a sphere at one-half

its distance from the source as at A, consequently the energy transmitted per square centimeter of surface will be only one-fourth as great at C as at A. The energy or intensity therefore varies inversely as the square of the distance from the source.

Of course if the wave is prevented from spreading out, as in a speaking tube or when the source of sound is near the surface of a smooth lake, this law does not hold.

312. Speaking Tubes and Ear Trumpets. The ordinary *speaking tubes* connecting distant rooms in buildings depend

Fig. 167

on the fact that the sound waves coming from the speaker, instead of spreading out in all directions from the mouth, are limited by the walls of the tube, so the sound may be heard with only moderately diminished intensity at the distant end. Bends and particularly sharp turns should as far as possible be avoided because reflections will then arise which may break up the advancing waves, reduce the intensity and interfere with the clearness of the sound.

When one end of a rod or wire of metal or of a long uniform beam of wood is struck the sound is carried along the rod or beam just as in a speaking tube with very little loss through waves sent out sidewise, and is therefore very distinctly heard at the farther end.

In ear trumpets, by the constraint of the smooth walls of the tube, the wave entering the wide end is gradually diminished in area, and although much of its energy may be lost through being reflected from the sides of the trumpet and back out of it again, *the energy per cubic centimeter* in the emergent wave is many times greater than that in the wave which entered the trumpet.

313. Megaphone and Speaking Trumpet. In the megaphone sound waves coming from the speaker instead of spreading out in all directions from the mouth are limited by the walls of the instrument, so that the wave emerging at the wide end has the whole energy of the voice. It will be shown in connection with
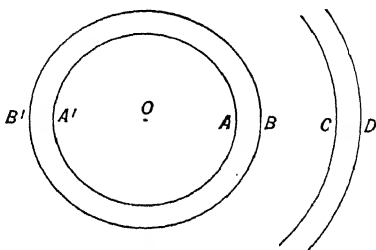
the diffraction of light that when light waves pass through an opening which is not more than a wave length in diameter, the waves spread out in every direction from the opening, while if the opening is much larger, the waves, on account of interference, will
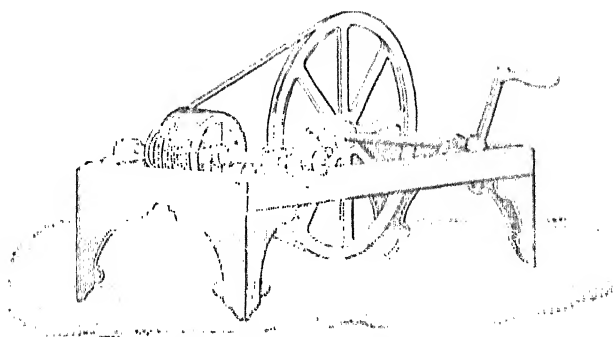
FIG. 168.   Savart's wheel

not spread out so much but will travel straight forward, illuminating a spot directly opposite the opening.   For the same reason sound waves coming directly from the mouth spread

FIG. 169.   Siren

out in every direction while waves from the larger opening do not spread out so much, and produce a more intense effect directly ahead.
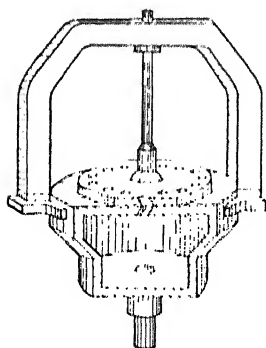
Exactly how this effect is caused by the interference of waves will be better understood after studying the diffraction of light.

314. Pitch. *The pitch of a sound depends on the frequency of the vibrations.* This is well shown by Savart's wheel. If a card is held so that it is struck by the teeth of a rotating cogged wheel a sound is given out which rises steadily in pitch as the speed of the wheel increases.   If a device indicating the number of revolutions is attached to the wheel the number of taps per second producing a sound of a given pitch is readily determined.

Another instrument by which the number of vibrations may be determined is the *siren* devised by Cagniard de la Tour, shown in

figure 169. A disc having a circular row of equidistant holes is mounted on an axis so that it can rotate almost in contact with the upper surface of a flat circular box in which holes are made exactly corresponding to those in the disc, so that as the disc rotates the holes are alternately opened and closed as many times in each revolution as there are holes in the series. The box is connected by a tube with a bellows and the puffs of air that come through the holes of the disc as it rotates give rise to a tone which
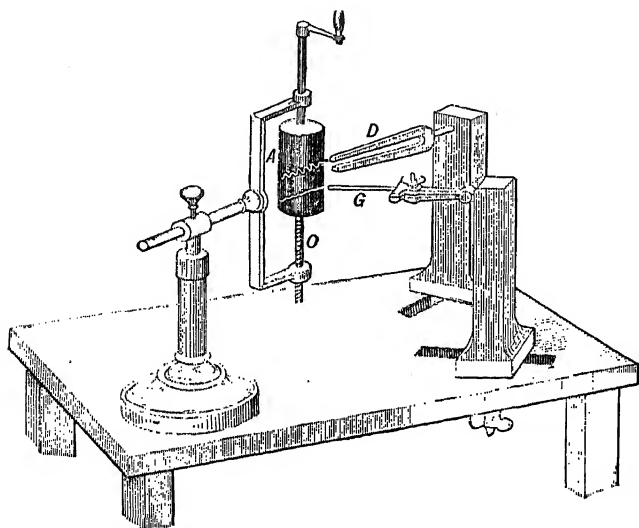
FIG. 170. Vibrations recorded on blackened drum

is higher in pitch the faster the disc rotates. A revolution counter is attached to the axle so that the speed may easily be determined.

The holes of the disc are inclined one way and those in the upper plate of the box are oppositely inclined, so that the blast of air through the holes causes the rotation to take place automatically, the speed being controlled by the strength of the blast and a brake if necessary.

For finding the number of vibrations of a tuning-fork *the graphic method* may be used, illustrated in figure 170. A point or stylus is fixed to one prong of a tuning-fork which is mounted so

that the stylus just touches a sheet of smoked paper stretched over a cylindrical drum. The axle of the drum is a coarse screw by which the drum is moved slowly lengthwise as it rotates. If the fork is set vibrating, on rotating the drum a wavy curve will be drawn in helical form around the drum, each wave corresponding to a vibration of the fork. To find the number per second a second curve may be simultaneously drawn alongside of the first by a tuning-fork whose frequency of vibration is known; or a small electric marker connected with a clock may be mounted with its point touching the drum close beside the stylus of the fork, so that its marks made every second lie close to the curve drawn by the fork. The number of vibrations of the fork per second are found by counting the undulations in the curve between two consecutive time marks.

These various methods of experiment show that the pitch of a sound *depends only on the frequency of vibration*, and that it
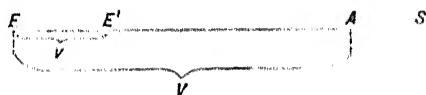
Fig. 171. Case of ear moving toward a sounding body

makes no difference whether the sound comes from a tuning fork or from the puffs of a siren or the taps of Savart's wheel, all will have the same pitch if the frequency is the same.

315. Doppler's Principle. The pitch of a sound *as heard depends on the number of waves that reach the ear per second*. Consequently if the ear is moving toward the sounding body the apparent pitch will be raised, since more waves per second will meet the ear; and conversely if the ear is moving away from the sounding body it will receive fewer waves per second than if it were at rest and the pitch will appear lower.

Let E (Fig. 171) represent the position of the ear and S that of a sounding body making n vibrations per second, and let the distance from A to E be V, the distance that sound travels per second. Then between A and E there are n sound waves which will reach the ear in one second if it remains at E, but if in one second the ear advances a distance v to E' it will meet in addition the waves between E and E'. Let x be the number of waves

between E and E', and n' the number reaching the ear per second, then

$$n' = n + x$$

and

$$x : n = v : V \quad \therefore \quad x = \frac{nv}{V}$$

and

$$n' = n \pm \frac{nv}{V} = n\left(1 \pm \frac{v}{V}\right)$$

the signs being plus or minus according as the ear has a velocity v toward, or away from, the sounding body.

A similar change in pitch is observed when the sounding body is moving toward or away from the observer. But in this case the formula is somewhat different as the wave length of the sound is changed in consequence of the motion.

Let S, the source of sound, have a velocity v toward the observer at E. In one second as it advances from S to S' it gives out n waves. The first of these waves leaving it at S has reached A, having advanced a distance V equal to the velocity of sound in the medium, by the time that the sounding body giving out the nth wave has reached S'. All n waves, therefore, lie between A and S', and the wave length, λ', is $\dfrac{V - v}{n}$.

But the number of waves that will reach E per second will be the number of wave lengths that are contained in the distance that the waves travel per second, or $n' = \dfrac{V}{\lambda'}$;

hence

$$n' = \frac{nV}{V \pm v},$$

where the minus sign is to be taken when the velocity of the sounding body is toward the hearer.

Fig. 172. Moving source of waves

Doppler's principle explains the sudden lowering in pitch observed in a locomotive whistle as it passes. It has also a most interesting application to light waves (§ 942).

316. Quality of Sound. The ear readily observes the difference in *quality* or *timbre* between the sound of a violin and that of a flute or between notes of an organ and those of a piano, though of the same pitch.

This individual character of tones depends in part on certain superficial characteristics. The note of the piano comes impulsively, suddenly strong, and then rapidly dying out, while the tones of an organ do not come instantly to full strength, but are then sustained and steady. But tones equally sustained and steady may yet differ greatly in *quality*, as, for example, the tones of tuning-fork, organ pipe, and violin. To investigate the cause of this difference we shall need *resonators* which vibrate in sympathy with the tones studied.

317. Sympathetic Vibration. A bell ringer by timing his pulls on the rope to correspond to the swing of the bell is able to set a heavy bell strongly swinging, while mere random pulls would accomplish very little; so it is that sound waves or other comparatively slight impulses may set up strong vibrations in a body if they are exactly timed to correspond to its natural period of vibration. This fact of *sympathetic vibration* may be illustrated by tuning two strings on a sonometer to the same pitch, and then sounding one strongly; the other will be set in vibration by the impulses communicated to it through the supporting bridges.

Again, if the dampers are raised from the strings of a piano and a clear strong note is sung near the instrument, the corresponding string will be heard sounding after the singer's voice is silent.

A very interesting case of sympathetic resonance is that in which a tuning-fork is set in vibration by the sound waves from a similar fork placed 20 or 30 ft. away. The two forks are mounted on suitable resonance boxes and must be of exactly the same pitch; if they are thrown out of unison even very slightly, as may be done by affixing a bit of beeswax to a prong of one of them, they will no longer respond to each other.

318. Resonators. When water is poured into a tall cylindrical jar the noise produced has a noticeable *pitch* which grows higher

as the water level rises in the jar. This pitch is due to the air column in the jar, which has a natural time of vibration of its own and responds to any component vibration of the same pitch which may exist in a noise produced in its vicinity. On blowing sharply across the mouth of the jar the same pitch is noticed, the confused rustling noise having some component to which the jar can respond. The roaring heard in sea shells is explained in the same way.

If an ordinary tuning-fork, not mounted on a resonance box, is held by its stem and struck, it will scarcely be heard a few feet away, but if it is held, as shown in the figure, over the mouth of a jar tuned to respond, a strong tone will be given out. The arrangement shown in the figure permits the tuning to be easily effected by raising or lowering the connected water reservoir, thus changing the level of the water in the resonance tube until the response is most powerful.
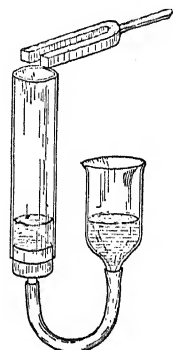
Fig. 173. Resonance

The pitch of the air column may be lowered also by partially closing the opening of the jar.

Such an air column being easily set in vibration by the proper tone is known as a *resonator* and may be useful in detecting in a mass of tones the presence of the particular one to which it is tuned.

319. Helmholtz Resonators. Helmholtz, in the analysis of composite tones, made use of spherical resonators, each having a large opening and also a small one adapted to the ear.

A resonator of this form is particularly useful because it *responds easily to vibrations of one pitch only* and so is well suited to the analysis of sound.

320. Complex and Simple Tones. The following experiment will now give us a clue to the cause of the difference in quality of tones. Take a series of tuning-forks mounted on resonating boxes, the frequencies of the forks being in the order of the series of whole numbers, 1, 2, 3, 4, 5, etc., which is known as the *harmonic* series. If the deepest toned fork in the series makes 250 vibrations per second, the next will make 500, and the next

750, etc. Provide also a set of Helmholtz resonators, one adapted to each fork.

Now, on sounding the lowest pitched fork alone, a deep tone is obtained to which *the corresponding resonator alone will respond*. If the next lowest is now sounded at the same time with the other, the tones blend and are heard as a single tone of the same pitch as before but of a different quality. And so by sounding along with the deepest or fundamental fork any or all of the others, making some of the component tones strong and some weak, great variety of tones may be obtained differing in quality though all are heard as of the same pitch.

But if any of these tones is tested by the resonators it is found that all those resonators respond which correspond to the forks used in producing the tone. Such a tone is called *complex*, while a tone to which only one resonator will respond is called a *simple tone*.

321. Analysis of Sounds. In the case just considered it is evident from the way in which the sounds of various qualities were produced that they were complex and consisted of sounds of different pitches blended together. But if we now sound an open organ pipe of the same pitch as the deepest toned resonator we find that not only does that resonator respond, but so also to a greater or less degree do the whole series of resonators, showing that though the sound comes from a single pipe it is just as truly complex as though originating in a series of tuning forks.

The component simple tones which unite to form a complex tone are known as its partial tones, the lowest of these in pitch is the fundamental, and the others are the upper partial tones or upper harmonics. The latter term is especially applicable when the upper partial tones are members of the harmonic series (§ 320) which starts with the fundamental.

From the laws of dynamics as well as from experiment there is reason to believe that *a simple tone, to which a resonator of only one certain pitch will respond is one in which the vibrations of the air are simple harmonic* (§ 124).

With attentive listening the ear may hear the simple harmonic components of a complex tone as separate simple tones. Persons with ears trained to the analysis of sound can often detect the different harmonics in a tone without the aid of resonators.

322. Synthesis of Sounds. Helmholtz devised an interesting apparatus by which complex sounds might be built up from their simple components. This consisted of a set of ten tuning-forks, corresponding to the first ten terms of a harmonic series, which were kept continuously vibrating by means of electro-magnets, each fork being mounted in front of an appropriate resonator as shown in figure 174.

The resonators were cylindrical brass boxes, each mounted with its opening close to the prongs of the corresponding fork,
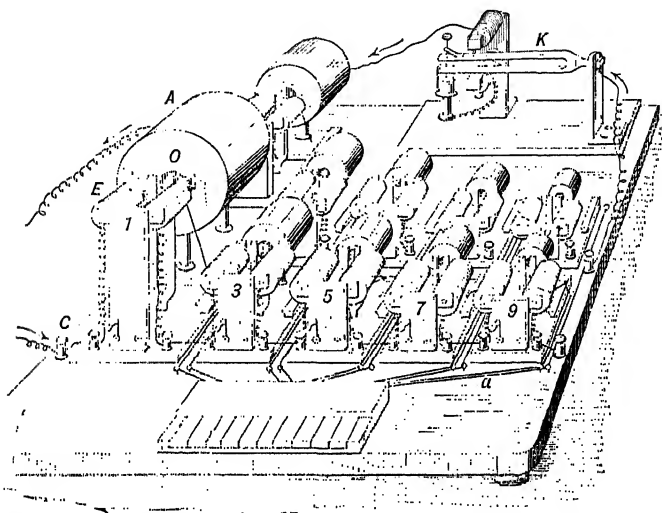
FIG. 174. Helmholtz apparatus for the synthesis of sound

the openings being closed by covers which could be drawn back by pressing the keys of the key-board. When the resonators were closed scarcely any sound came from the forks, but drawing back the cover from any resonator by depressing its key brought out the corresponding tone with an intensity which depended upon the amount that the key was depressed. By means of such an apparatus the sound of an open or closed organ pipe, a violin, or reed instrument can be closely imitated.

An interesting modern instance of the synthesis of sounds is found in the ingenious "telharmonium" of Mr. Cahill in which the separate harmonics are transmitted by means of alternating

currents of electricity of different frequencies which combine to form a single resultant current which acts on the telephone receiver at the end of the line. By combining the proper simple harmonic components all the instruments of an orchestra are imitated.

323. Quality of a Musical Tone. The quality of a musical tone may then be said to be determined by the pitch and intensity of the different simple tones or harmonics into which it may be resolved.

324. Fourier's Analysis. It was shown by the distinguished French mathematician, Fourier, that any regular periodic vibra-
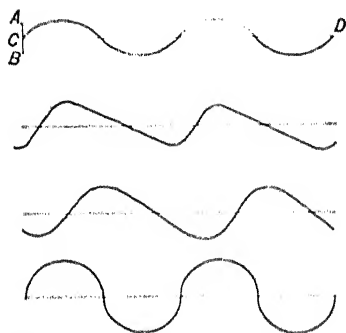
FIG. 175. Vibration curves   FIG. 176. Combination of vibrations
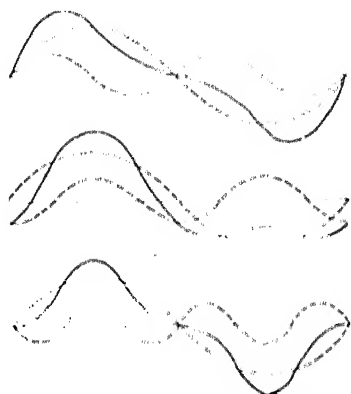
tion, such as can take place in a sound wave, may be resolved into a sum of simple harmonic components all of which belong to a harmonic series, in which the fundamental has the same period as the vibration analyzed. Thus, according to this theorem of Fourier, it is possible to analyze any sound wave into its simple harmonic components, and it is these simple harmonic components which are the simple partial tones detected by resonators.

For example, the upper curve in figure 175 represents a simple sine wave. The three lower curves represent waves having the same wave length and therefore the same periodicity as the upper curve, but they represent entirely different modes of vibration.

Now, according to Fourier's theorem, each of these curves can be resolved into simple harmonic components. In figure 176,

for example, the wave forms expressed by the heavy lines are the resultants of the simple harmonic waves represented by the dotted sine curves. The resultant curve in each case is obtained by adding the corresponding ordinates of the two component curves. In the first two cases one component has the same wave length as the resultant, while the other has half that wave length. These two components have the same amplitude in the first case as in the second, but their relative phases are different in the two cases and hence the resultant curves are different. In the lower curve one component has one-third the wave length of the resultant, while the component having half the wave length is absent or has zero amplitude.

325. Musical Tones. In tones suitable for music the upper partial tones fall almost exactly into the harmonic series, starting with the fundamental; that is, their frequencies are very nearly exact multiples of the frequency of the fundamental.

But the partial tones given out by bells and plates when struck do not correspond, even approximately, to the lower terms of the harmonic series, and are quite unsuited for music.
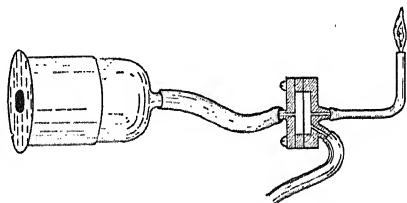
FIG. 177. Resonator and manometric flame

326. Koenig Resonators and Manometric Flames. The French acoustician, Koenig, made use of resonators in which the front part was cylindrical and could be pushed in or drawn out so that each could easily be adjusted in pitch.

To observe the vibrations of the resonators he employed *manometric flames*. A small, flat, disc-shaped box or capsule of wood was divided into two chambers by a thin membrane, such as gold beater's skin. The cavity on one side of the diaphragm was connected by a short tube with a resonator, while the cavity on the other side had two openings, through one of which illuminating gas was admitted, while the other was connected with a fine jet where the gas burned in a small flame. The vibrations of the air in the resonator were transmitted through the diaphragm in the manometric capsule to the illuminating gas, causing the flame to dance.

The image of such a flame viewed in a rotating mirror is drawn out in a band of light which when the flame is in oscillation shows serrations like saw teeth, as shown in figure 178; the particular form of the serrations revealing the mode of vibration of the flame.
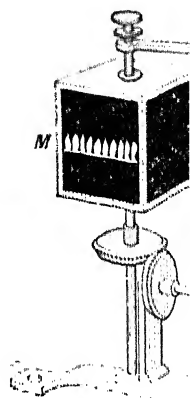
327. Sensitive Flames. Under some conditions a gas flame may be very sensitive to sound. A small cylindrical jet is required having an aperture about 0.5 mm. in diameter, and the pressure must be such as to produce a long flame *just on the brink of roaring*. A pressure of about 9 in. of water is commonly required. Such a flame is sensitive to the vibrations of exceedingly short waves of sound, and breaks into a shorter flaring or roaring flame when a bunch of keys is jingled in its vicinity or a sharp hiss given or a very high-pitched whistle sounded.

By means of sensitive flames sound waves so short as to be quite inaudible may be detected, and their interference and reflection studied.

328. Photography of the Shape of Sound Waves. The direct photography of wave forms of sound vibrations has been successs

Fig. 179. Phonodeik

fully accomplished by Professor Dayton C. Miller by means o sensitive instrument invented by him called the *phonodeik*. T principle of this instrument is shown in figure 179. Sound wav enter the horn *H* and produce vibrations of the delicate gl: diaphragm *D*. A tiny mirror *M* attached to a steel spindle set

jewelled bearings is made to rotate back and forth exactly with the diaphragm motions by means of two or three fibers of silk attached to the center of the diaphragm, and passed once around a small pulley on the steel spindle and fastened to a spring, S. A ray of light from a small hole is focussed by the lens L and is reflected from the mirror to the moving photographic film F which registers the motion of the spot of light. Figures 180 and 181
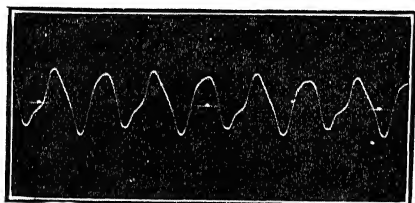
FIG. 180. Vowel sound *oo*

show records of the vowel sounds *oo* and *ee* obtained in this way. The film records give the diaphragm motions on a greatly magnified scale, so that it is possible to analyze the sound into its various harmonic components (§ 324) with considerable accuracy. By means of special groups of adjustable organ pipes Professor Miller has not only reproduced various musical tones
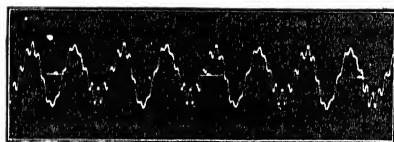
FIG. 181. Vowel sound *ee*

but also the sounds of the human voice by accurately superimposing the harmonics which these sounds were found to contain from a study of their film records.

329. Photography of Sound Waves in Air. It is a very interesting fact that sound waves produced by sudden concussions, such as those produced by a spark of a large Leyden jar, or by the discharge of a gun, can be photographed as the waves travel through the air. Figure 182 shows shadow photographs of sound waves spreading outwards from a knob where a spark discharge has just taken place. The photographic plate is
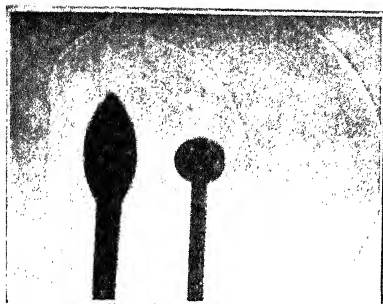
A. Carbon-dioxide lens. Spark at principal focus. Refracted wave plane
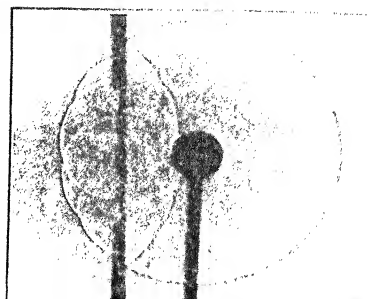
B. Plane grating, four apertures Wave .00016 second after spark. Huygen's Principle (§964) verified by both the reflected and the transmitted waves
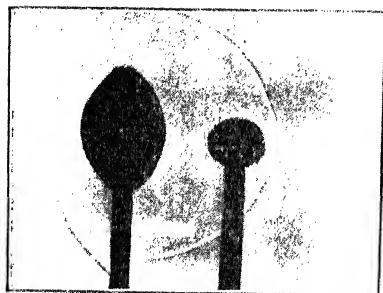
C. Hydrogen lens. Curvature of refracted wave slightly increased

D. Lens of sulphur dioxide gas. Refracted wave convergent

FIG. 182. Sound wave photographs. Sound wave starts at the knob at center. (Photographed by Professor A. L. Foley)

screened from the light of the spark, which produces the sound wave, but is illuminated by the light from a second spark produced a little later than the first spark at a greater distance from the photographic plate. The light from the second spark which passes through the condensation wave produced by the first spark has its path bent slightly. This results in an unevenness of illumination of the photographic plate, which on development reveals the presence of the sound wave as shown in figure 182. The photographs of figure 182 were taken by Professor A. L. Foley of Indiana University.

By this method sound waves from the muzzle of a gun and also the bullet with ripples produced by it in the atmosphere can be seen.

330. Architectural Acoustics. The photography of sound waves has also been employed by W. C. Sabine in the study of the acoustic properties of theaters and halls, to find out how echo effects may be reduced. The course of the sound waves and their echoes were observed in small cross section models of the theaters studied.

Sabine found that for the best hearing, halls should have echoes which last about one second. When the echoes were more prolonged they interfered with good hearing, and if they died out too soon the sound was too dead to be pleasing. The walls and interior parts were built of or partially covered with suitable sound absorbing material to cut the echoes down properly, which might be in the form of curtains or a special sound absorbing compound. Ordinary wood and masonry reflect so much sound that they are objectionable as surface materials of large rooms.

The shape of a hall or theater is important, since the sound may come to a strong focus at certain spots. This effect is illustrated in the famous whispering gallery in the dome of St. Paul's Cathedral in London where a low whisper on one side is clearly heard on the other, and sounds as though the source were close to the listener.

REFERENCE

D. C. MILLER: *The Science of Musical Sounds* (MacMillan).

INTERFERENCE AND BEATS

331. Superposition of Waves. When the same portion of fluid is traversed by two waves, the motion of the particle will be the resultant of the two and may thus become very complicated.

In case of surface waves in a liquid the eye can readily observe a series of short waves running over longer ones and preserving their motion as though over an undisturbed medium.

332. Interference of Waves. Suppose A and B are the centers of two series of waves of the same wave length and amplitude. Then there will be certain points, as at C, where waves from the two sources act together so as to produce great disturbance. Suppose the lines in the diagram represent the crests of waves, then at the points C the crest of one wave is superposed on another, while at C'' the troughs of two waves come together; a half period later the crests will

be at C'' and the troughs at C. There will result, therefore, along the line CC'' a series of waves of double the amplitude of the original waves. At the points D, however, the crest of a wave from one center coincides with the trough of a wave from the other, therefore there will be at least a partial neutralization at points along the line DD'. If the waves coming together at D have equal

FIG. 183. Interference of waves

amplitudes and equal wave lengths and are also simple harmonic waves they will separately produce at D equal and opposite displacements at every instant and will therefore completely neutralize each other.

This interaction of two sets of waves by which at certain points one is more or less completely neutralized by the other is known as *interference*.

333. Energy is Not Lost in Interference. When there is complete interference at any point there is no motion of the medium and no energy at that point, but the energy of the two interfering waves is not lost or destroyed but appears at neighboring points (such as C, Fig. 183) where the amplitude of the component waves is added. For at these points the energy of the resultant vibration is four times what it would be if one of the trains of waves were suppressed. There results from the interaction of the two wave systems *a different distribution* of energy, but the total energy remains unchanged.

334. Interference of Sound Waves. The interference of sound waves is well shown in the following experiment. The

sound waves from a tuning-fork (Fig. 184) enter a suitable receiver which is connected to an ear-piece by means of two tubes, one of which has a sliding portion by which its length can be varied. When the tubes are adjusted to be of equal length the sound of the fork is distinctly heard by the observer at *E*. As the sliding tube is drawn out, making one tube longer than the other, the sound grows fainter and reaches a minimum when one tube is longer than the other by a half wave length of the sound waves sent out by the fork, for in this case the waves reach the ear through the two tubes in opposite phases and interfere. If the slider

FIG. 184

is drawn out still farther the sound increases in strength reaching a maximum when one tube is just a whole wave length longer than the other.
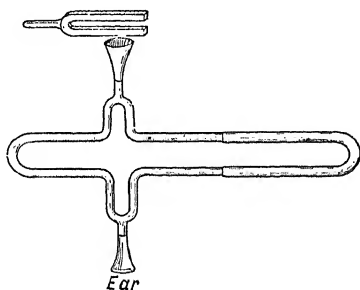
If two similar organ pipes of the same pitch are mounted on a rather small air chest, as shown in figure 185, and sounded

FIG. 185. Interference between organ pipes

simultaneously they will usually sound in opposite phases, owing to an oscillation of the air in the air chest itself. The sound waves coming from one pipe will thus interfere with those from the other and the *fundamental* tones will be almost completely neutralized, the higher harmonics will, however, still be heard.

335. Beats. If two organ pipes sounding together are not exactly of the same pitch the sound comes in pulses or throbs called *beats*. For in this case one pipe is giving out more vibrations per second than the other and, consequently, the relative phases of the two are constantly changing, as shown in the following figure where the dotted curves represent the waves from the two pipes, one of which is supposed to give out eleven vibrations for every ten of the other. The full line represents the resultant motion. It is clear that while one pipe is gaining one complete vibration on

the other, there will be an instant when the waves are in opposite phase and interfere and another instant when they will be in the same phase and strengthen each other. There will, therefore, be one hundred beats in the time in which one pipe has made one hundred more vibrations than the other. Or *if one pipe makes m vibrations per second and the other n, the number of beats per second is m − n.*

Beats are easily heard when two adjoining notes on the piano or organ are simultaneously struck, and the lower the notes are on the scale the slower will be the beats. Beats are not heard, however, between notes that are very different in pitch. For example, no beats would be heard in case of two simple tones, one making 200 vibrations per second and the other 300 vibra-
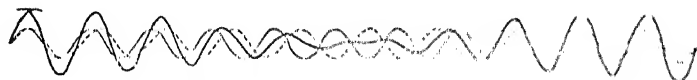
FIG. 186. Formation of beats

tions, though if one made 2000 vibrations and the other 2100 there would be 100 beats per second, heard as a distinctly jarring roughness. The explanation of this was given by Helmholtz (§ 371).

In tuning two strings or two forks to unison they are adjusted until no beats are heard. If it is required that two tuning forks shall be very accurately of the same pitch they may be tuned to make the same number of beats per second with a third fork, as it is easier to count accurately three or four beats per second than to distinguish between no beats at all and very slow beating.

STANDING WAVES AND VIBRATING BODIES

336. Transverse Vibration of a Cord. Take a long flexible rubber tube or other elastic cord fixed at one end, as at *P* (Fig. 187), and, holding it slightly taut, let the hand give a sudden movement to one side and back again.

A wave is set up as at *A* which runs the length of the cord, is reflected at *P*, and returns on the opposite side of the cord as shown at *B*. On reaching the hand it is reflected to *A* and the motion is repeated. *Thus the cord makes a complete vibration*

and returns to its original form in the time in which a wave runs the length of the cord and returns.

If the wave is not sent out by so sharp a movement it may take the form shown in the lower part of the figure where the cord simply swings to and fro or vibrates sidewise.

But this vibration is just as truly due to a wave running along the cord and back again as in the first case and *the period of one complete vibration is the time required for a transverse wave to run the length of the cord and return.*

Now, suppose waves are sent out by the hand with *twice* the former frequency, a wave will start from the hand at the same instant that the preceding wave starts back after reflection. The two will meet at the middle and, as they are on opposite sides of the cord in consequence of reflection, they will exactly neutralize each other at the middle point, which will, therefore, remain at rest and the cord will vibrate in two segments. The vibrating segments are called *loops* and the points of rest *nodes*. The period of a complete vibration is in this case only half that of the fundamental period when there is only one loop.
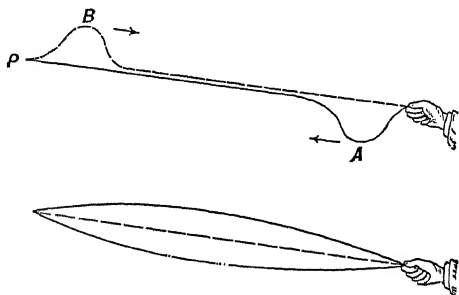
FIG. 187.   Waves in cord

If waves are sent out with three times the frequency of the first case, the cord will break up into three vibrating segments, with intermediate nodes, and with greater frequency of vibration a still greater number of segments may be produced.

337. Standing Waves. These vibrating segments are a particular case of what are called *standing waves, which are set up in water or air or in other elastic bodies by the interaction of similar trains of waves running in opposite directions*, and are usually due to reflected waves meeting those which are advancing.

Standing waves are easily observed on the surface of water in a circular vessel in the center of which a periodic disturbance is produced. If the period of the disturbance is so adjusted

that the wave length produced has the proper relation to the
size of the vessel, a steady state is produced in which waves going
outward meet the reflected waves, causing nodal rings where
the water is at rest. Between these rings the surface oscillates
up and down.

Standing waves are also produced in organ pipes by the re-
flection of the air waves from the ends of the pipes.

338. Formation of Standing Waves in Cord. The diagram
(Fig. 188) illustrates the mode of formation of nodes and loops in
a cord.

The dotted line represents a wave traveling from right to left
along the cord, while the broken line represents an equal train
of waves moving from left to right as indicated by the arrows.
The resultant wave is shown by the continuous curved line,
and its ordinate at any point is the sum of the ordinates of

Fig. 188.   Formation of nodes and loops

the two component curves at that point. It will be observed
that the crests of the two component waves are approaching each
other at the points marked *L*, and a moment later will coincide.
The resultant wave will then be at a maximum. A quarter
period later the two component waves will exactly neutralize
each other, the crest of one coming exactly over the trough of the
other, and the cord which takes the resultant form will at that
instant be straight. As the waves move still farther the crests
A and *C* will come together in the middle of the diagram and the
resultant wave will then show a form just opposite to that in
the figure, being bent up in the middle and down on each side.
A little consideration will show that there will never be any dis-
placement at the points *N*, *N'*; for in any position of the two
component waves one is always as much above such a point as
the other is below it; these points are therefore nodes, and *the
distance between consecutive nodes is one-half the complete wave
length.*

339. Velocity of Wave in a Cord. The velocity of a transverse wave along a stretched cord may be deduced as follows. Suppose that an infinitely long cord having tension T and a mass per unit length m is drawn rapidly through a bent glass tube as shown in the figure. If the cord were at rest it would produce a pressure against the tube in consequence of its tension. At a point where the radius of curvature of the tube is r the pressure against the tube, or force per unit length, is $\dfrac{T}{r}$, being greater the more

Fɪɢ. 189. Solitary wave in cord

sharply curved the tube is at that point. But if the cord is drawn through the tube with velocity V, its centrifugal force per unit length as it runs over the curved part of the tube is $\dfrac{mV^2}{r}$, and this acts to diminish the pressure. If the speed is just right, one will exactly balance the other and we shall have

$$\frac{mV^2}{r} = \frac{T}{r} \quad \text{or} \quad V^2 = \frac{T}{m}.$$

Since the radius of curvature cancels out of the result, the speed at which there will be no pressure against the tube will be the same whatever its radius of curvature may be, and consequently whatever its shape. Suppose the cord is now drawn along at this critical speed, the tube may be made to vanish and the bend in the cord will remain unchanged. If, now, the observer moves along at the same rate as the cord from left to right, the cord will appear to him to be at rest while the bend will be seen to travel along the cord as a wave from right to left with a velocity V, where

$$V$$

The velocity with which a transverse wave runs along a perfectly flexible cord is thus determined from the relation

Velocity of the wave in cm./sec. $\quad \sqrt{\dfrac{\text{Tension in dynes}}{\text{Mass per cm. in grams}}}.$

340. Transverse Vibration of Cord. It has already been shown that the time of vibration of a stretched cord when it vibrates as a whole is the time required for the wave to run from one end of the cord to the other and back again. Or if it is vibrating in segments the period of vibration in one of the

segments is the time required for the wave to run twice the length of the segment. If l be the length of the cord or the distance between consecutive nodes, the period of vibration P is

$$P = \frac{2l}{V}$$

and if n is the frequency of the cord, or number of vibrations per second, we have

$$n = \frac{1}{P} = \frac{V}{2l}$$

where V is the velocity of a transverse wave along the cord. Substituting the value of V from the preceding paragraph we have

$$n = \frac{1}{2l}\sqrt{\frac{T}{m}},$$

a formula which expresses in compact form the following three laws:

1. The number of vibrations per second made by a string under a given tension is inversely proportional to the length of the vibrating segment.

2. In case of two strings of the same length and mass per unit length, the frequencies are proportional to the square roots of the tensions; thus if one has four times the tension of the other it will make twice as many vibrations per second.

3. If two strings have the same length and are under equal tensions, their frequencies will be inversely proportional to the square roots of their masses per unit length. Thus if one is four times as heavy as the other, it will have only half the frequency.

341. Upper Harmonics of Cord. It has already been seen that a stretched cord may vibrate not only as a whole, but it may also vibrate in two segments or in three, and so on. These modes of vibration may be easily established by touching the cord lightly at a point where a node is desired and at the same time bowing it at a loop. For instance, if a cord is touched at one-fourth of its length from one end and then bowed half way between the end and the point where it is touched it will vibrate in four segments and give a tone which has a frequency four

times that of its fundamental mode of vibration. That the cord actually vibrates in this way was prettily shown by Tyndall as follows: Little riders of bent paper were hung on the cord at the points where nodes were to be established and others midway between them on the loops. On sounding the cord as above described all the little riders on the loops were "unhorsed," while those at the nodes remained undisturbed.

These partial modes of vibration of a cord are known as its *harmonics*, because if the fundamental mode of vibration of the cord has a frequency of n vibrations per second the partial modes of vibration will have frequencies $2n$, $3n$, $4n$, $5n$, etc., according as the cord vibrates in 2, 3, 4, or 5 segments, respectively; thus the frequencies of the partial modes of vibration are related to the fundamental as the terms in a harmonic series.

If cords were *perfectly flexible* this would be *exactly* the case, but as actual cords always have a certain amount of stiffness, which affects the higher harmonics where the vibrating segments are short more than the lower harmonics where the vibrating segments are longer, the above is simply a close approximation to the truth, and the cord when vibrating in four segments, for example, will have slightly *more* than four times the frequency of the fundamental.

342. Superposition of Vibrations. When a string is struck or bowed, a number of these different possible modes of vibration
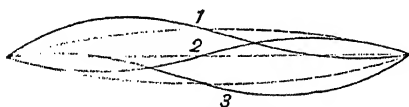
FIG. 190. Two simple vibrations combined

are in general set up *simultaneously*. The form of vibration assumed by a cord in which two such modes of vibration coexist is shown in figure 190. The dotted lines show three positions of the cord simply vibrating as a *whole* in its fundamental mode. The full lines, however, show its form when it is at the same time vibrating in two segments, the two halves vibrating with reference to the dotted lines just as they would have done about the middle straight line if the fundamental mode of vibration had been absent.

343. Experimental Demonstration. That such a coexistence of different modes of vibration commonly exists in a string may be shown by the following experiment. Pluck the string of a sonometer strongly at, say, one-fourth of its length from one end; then touching the string lightly at the middle, its fundamental mode of vibration will be damped, but it will still be free to vibrate in two segments and the tone characteristic of that mode of vibration, an octave higher than the fundamental, will be heard still sounding. Or if struck one-sixth of its length from one end and then touched lightly at one-third of its length, the fundamental mode will be damped while it will still be heard sounding a tone having three times the frequency of the fundamental.

344. Young's Law. *When a string is struck at any point only those modes of vibration are set up which do not have nodes at that point.*

When a string is touched at any point all vibrations are damped that do not have nodes at that point.

It is clear from these laws that in order to stop all the vibrations of string it should be damped at the same point where it is struck. This is done in the piano.

345. Quality of Tone. Strings are suited for musical instruments because the partial tones that they give out have frequencies which form a harmonic series with the fundamental, and all the lower tones of a harmonic series as far as the seventh harmonic are pleasing when sounded together. The seventh and ninth harmonics are, however, decidedly inharmonious with the others, and therefore it is desirable that in musical instruments the strings should be struck or bowed in such a way that these harmonics may not be developed. This is accomplished in the piano by striking the strings about one-seventh or one-eighth of their length from one end, so that all the partial modes of vibration having nodes near that point are weak.

The hardness of the hammer in a piano also has a decided influence on the tone. The harder the hammer the more sharply the string is bent when struck and the more prominent are the higher harmonics. If the hammer is too soft the tone is soft and lacking in the richness that comes from the proper strength of the harmonics.

Organ Pipes and Wind Instruments

346. Organ Pipes. An organ pipe may be considered as made up of two parts — a vibrator and a resonator. There are two types in use, *flute* pipes, in which the vibrations are caused by a stream of air rushing against an edge, and *reed* pipes, in which the vibrator is a thin strip of metal.

The construction of a *flute* pipe is shown in figure 191. At the lower end of the pipe is the embouchure or mouth which is like that of an ordinary whistle. Air, forced into the air chamber at the bottom, escapes through a narrow slit against an edge just opposite. The upper part of the pipe is a tube which may be either open or stopped at the upper end, and constitutes a resonator which reinforces the vibrations set up at the embouchure.

Fig. 191. Section of organ pipe

If a blast of air is sent through a skeleton pipe, which has a mouth-piece but no resonating chamber, a soft whistling noise is heard which rises in pitch with the force of the blast. The pitch of this tone also depends on the bluntness of the edge against which the blast strikes and its distance from the opening. If the pipe is now provided with a resonating chamber of a proper shape and size to reinforce the vibration, a strong, clear tone will be given out.

The vibrations appear to be caused by the friction of the stream of air against the edge, together with its inertia, just as little waves are formed on the surface of a stream of water in front of a wire or rod which cuts the surface.

The vibrations caused in this way are taken up by the air column in the pipe, which as it vibrates reacts on the stream of air at the mouth-piece, causing it to be deflected alternately inward and outward in rhythm with the vibration of the air column; in this way regular impulses are received by the air column, and a strong vibration is maintained.

In case, then, of *flute* pipes it appears that *the size and shape of the resonating cavity is what chiefly determines pitch*, though a certain adaptation in the form of mouth-piece and strength of blast is required in order to evoke a good tone.

347. Nodes and Loops in Organ Pipes. The vibration of the air column in an organ pipe is a case of standing waves and is due to the interaction between waves running up the pipe and reflected waves moving in the opposite direction, forming nodes and loops just as in the case of the vibrations of a cord.

The upper part of figure 192 represents sound waves advancing from left to right in the direction of the upper large arrow. In condensed portions of the waves at C_1 and C_2 the particles are moving forward in the direction of advancement of the wave. In the rarefied portions at R_1 and R_2 the particles have a backward velocity. Immediately under this is represented the state of things in waves returning from right to left. The particles in the condensed portions of these waves have a velocity from right to left as shown by the small arrows, and from left to right in the rarefied regions. If, now, these two sets of waves pass simultaneously through the same mass of gas the particles take the resultant motion and nodes and loops are formed in
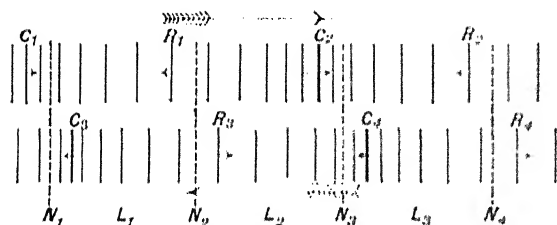
Fig. 192.   Formation of nodes and loops in organ pipe

the positions indicated in the lower part of the diagram. For it is clear that the rarefactions R_1 and R_3 will reach N_2 simultaneously, tending to produce opposite displacements, and a half period later the condensations C_4 and C_1 will come together at the same point, tending also to produce opposite displacements, and a little consideration will show that at every instant the two sets of waves will balance each other at N_2 so that the particles there will remain at rest. So also at N_1, N_3, and N_4, which are thus *nodes*. But at L_2 the rarefaction R_1 of the advancing wave will arrive at the same instant as the condensation C_4 of the returning wave, and since in both of these the velocity of the particles is from right to left the resultant velocity at L_2 will be from right to left at that instant. A half period later when C_1 and R_4 come together at L_2 the particles there will have a maximum velocity from left to right.

Thus the air between nodes surges back and forth, in one half vibration swinging toward N_2 on both sides and producing a compression there as shown in the lower part of figure 193. While in the next half vibration the air layers swing away from N_2 and toward N_1 and N_3, producing rarefaction at N_2 and condensations at N_1 and N_3, as in the upper part of figure 193.

At nodes, therefore, the greatest changes in pressure take place, although

the nodal layer itself remains at rest, while the motion of the particles is greatest in the loops midway between nodes.

Successive nodes are a half wave length apart and are in opposite phases, one being a point of rarefaction at the instant when the other is a point of condensation. So also *the phases of motion in successive loops are opposite.*
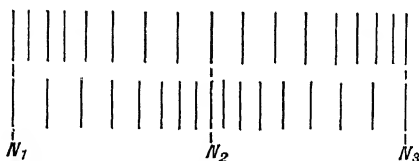
FIG. 193.   Opposite phases at nodes

348. Kundt's Experiment. The nodes and loops in a vibrating column of gas are beautifully shown in the following experiment due to Kundt.

A glass tube about 5 cm. in diameter and a meter long is tightly stopped at one end, while in the other is fitted a light piston of cardboard attached to the end of a glass rod which is clamped firmly at the middle. The rod is set in longitudinal vibration by drawing along it a wet cloth held firmly clasped around the rod.
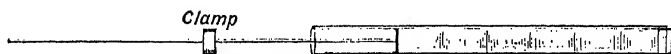
FIG. 194.   Kundt's tube experiment

By adjusting the position of the piston A in the large tube a point is found where the air column between A and B is in resonance with the vibrations of the rod. The air is then set in such powerful vibration that any light dust in the tube, such as lycopodium powder, is driven out of the loops and gathers in little heaps in the nodes. This will occur when the sound waves run the length of the air column and back in a certain whole number of vibrations of the rod. In the diagram there are six loops indicating that the rod makes six vibrations while the wave runs the length of the tube and returns.

The stopped end is exactly a node, while the piston end, where the motion is communicated, is *very nearly* a node.

The distance between nodes is a half wave length and may easily be measured with considerable accuracy. The tube may now be filled with some other gas and the distance AB again

adjusted and the distance between nodes found for this gas also, and since the frequency of vibration of the glass rod is the same in both cases the velocity of sound in the gas is to that in air in the same ratio as the distances between nodes in the two cases.

In this way the velocity of sound has been measured in a large number of gases and vapors.

349. Reflection in Organ Pipes. Reflection of waves takes place in general when a wave meets a boundary where there is a change of medium.

In a stopped organ pipe, a wave running up the pipe meets the unyielding end, and must therefore be reflected in such a way that
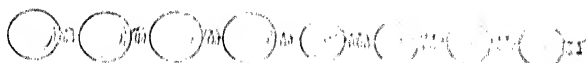
Fig. 195.   Reflection model

the reflecting surface is a *node*, or point of no motion. In an open pipe it is quite the reverse, the wave advancing in the pipe on coming to the open end finds a freer medium, unconstrained by the walls of the pipe, and therefore reflection takes place, but in such a way that the end is a *loop* or point of great motion.

The mechanical model represented in figure 195 will serve to make clear the nature of these two kinds of reflection. The figure represents a series of balls resting in a frictionless groove. Half are larger and of greater mass than the other half, and all are connected together by springs of equal stiffness. Suppose the left-hand ball is given an impulse forward, the spring between it and the next will be compressed and the motion transmitted as a wave of compression from one to the other and so on along the line, each coming to rest after giving up its motion to those ahead. But the last of the row of large balls will move forward more freely than it would have done if there had been no change in the size of the balls, and, therefore, it stretches the spring behind it, which gives a forward pull to the ball next behind it and so sends back a wave of rarefaction. This is the kind of reflection which takes place at the *open* end of an organ pipe.

When, on the other hand, a compressional wave is sent from right to left along the row of small balls, the last one does not

move forward as far as it would have done if there had been no change of medium, and the spring behind it is, therefore, more compressed and gives a backward impulse to the preceding ball, sending a compressional wave back through the series. This is the kind of reflection which takes place at the *stopped* end of an organ pipe.

350. Open Pipes. *In open organ pipes there must, therefore, be a loop at the top and also a loop at the mouth* for there is great motion of the air at these points. Between them near the middle of the pipe is a node.

The position of the node may be demonstrated by lowering into the pipe a horizontal tray of thin membrane covered with sand. If the tray is exactly in the node the tone is not affected, but if it is either raised or lowered a loud buzzing is heard in consequence of the vibration of the membrane.

An open pipe sounding in this way is giving out its deepest or fundamental tone. Since consecutive loops are a half wave length apart, it follows that *the length of an open pipe is half the wave length of its fundamental tone.*

Thus an open pipe 1 meter long gives out waves 2 meters long and accordingly makes $\dfrac{331}{2} = 165.5$ vibrations per second, since $n = \dfrac{V}{\lambda}$.

But an open pipe may vibrate in other ways consistent with the condition that the two ends must be loops. Thus it may vibrate having loops and nodes, as shown in figure 196, so that the distance from loop to loop or node to node may be only one-half the length of the pipe, or it may be one-third the length, etc. The frequency of vibration in the first of these partial modes is, therefore, twice that of the fundamental, that of the next three times, the next four times, etc.

Thus, *any of the harmonics of the fundamental may be produced by an open organ pipe.*

The partial modes of vibration generally coexist to a greater or less degree with the fundamental, giving character and richness to the tone. The relative strength of the harmonics depends on the shape of the mouth-piece and the force with which it is blown and the shape of the pipe itself. The higher harmonics are more emphasized when the pipe is strongly blown.

The position of the nodes and loops just discussed is only approximately correct, as the open end is not exactly a loop, still less is the mouth exactly at a loop and the node is nearer the mouth of the pipe than it is to the top, the variation being most marked in wide pipes.

351. Stopped Pipes. In stopped pipes the stopped end is exactly a node, while the mouth is nearly a loop. Thus *the length of the pipe is one-fourth the wave length of the fundamental tone.*

Suppose a compressional wave starts at the mouth of the pipe, it runs up to the top, is there reflected back as a compressional
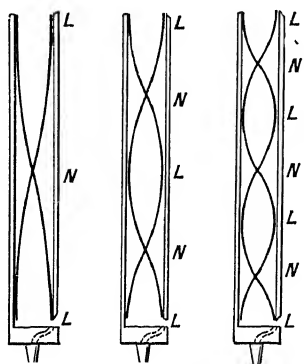
FIG. 196.   Nodes and loops in open pipes
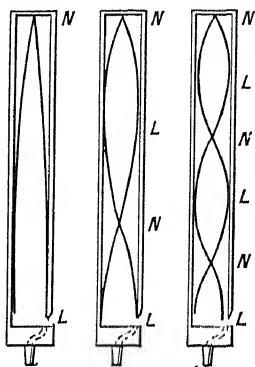
FIG. 197.   Nodes and loops in stopped organ pipes

wave, but on reaching the mouth again is reflected in the opposite way as a rarefaction, then traveling up and being reflected as a rarefaction it returns to the mouth where it is again reflected as a condensation; the wave has thus traversed the length of the pipe four times before returning to its original phase. A stopped pipe one-half meter long would therefore have a wave length two meters long and would vibrate with the same frequency as an open pipe one meter long.

Stopped pipes also may vibrate in other modes than the fundamental, but there must in every. case be a node at the top of the pipe and a loop at the mouth. In figure 197 is shown the distribution of nodes and loops in the fundamental and next two upper partial modes of vibration. It will be observed that the distance from node to loop in the first partial mode is one-third that in the

fundamental and consequently the frequency of this mode of vibration is three times the fundamental. The next higher mode of vibration has five times the frequency of the fundamental, etc. *Hence a stopped organ pipe may sound its fundamental and odd harmonics*, the frequencies of its proper tones being related to each other as the series 1, 3, 5, 7, etc. The absence of the even harmonics causes the tones of stopped pipes to differ decidedly in quality from those of open ones.

352. Reed Pipes. In reed pipes the vibrations are caused by a metal tongue or reed. Two forms are used, the *free* reed and the *striking* reed. A *free reed*, represented in figure 198, consists of a tongue of thin metal riveted firmly at one end to a plate
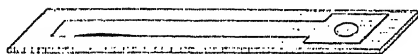
FIG. 198. Free reed

in which there is an aperture just under the tongue large enough to admit of its vibrating freely through it without touching.

The tongue when at rest is slightly above the aperture and when the reed is blown the stream of air catches it and carries it down, nearly closing the opening, this stops the rush of air and the tongue springs back, when the action is repeated. In this way a vibration is maintained.

Reeds of this type are used in cabinet organs, harmonicas, and accordions.

In *striking reeds* the tongue is a little larger than the opening and when at rest stands slightly above it. When blown it is carried down and clapping over the opening stops the rush of air, then rebounding is again carried down, thus being maintained in vibration. The tones from striking reeds are stronger and more penetrating than from free reeds. They are used in ordinary tin horns and in the clarinet and in some stops of pipe organs.

When reeds are used in pipe organs they are provided with resonators which strengthen and improve the tones.

The tones of reed pipes are rich in the higher harmonics, and the shape of the resonator used greatly influences the relative strength of these harmonics and hence determines the quality of the tone produced.

353. Effect of Changes of Temperature. When the temperature rises the velocity of sound in air increases and consequently the pitch of the flute pipes in an organ is raised. On the other hand, the effect of higher temperature is to diminish the elasticity of the metal tongues of reeds so that they vibrate more slowly, lowering the pitch of the reed pipes. An organ is thus thrown out of tune by great change of temperature.

354. Other Musical Instruments. In the flute and piccolo the vibrations are produced by blowing across an opening or embouchure near one end, the pitch produced being determined by the strength of blast and by the effective length of the resonating cavity which is regulated by opening or closing holes in its side. The deepest tone of a flute, as of an open organ pipe, is one whose wave length is double the length of the instrument. When blown hard the higher harmonics are sounded.

The mouth-piece of a fife is like an ordinary whistle or flute organ pipe, while the clarinet has a mouth-piece in which a thin slip of wood mounted over an opening forms a striking reed.

In the bugle the vibrations are due to the air being blown between the tightly drawn lips of the player as they are placed upon a suitable cup-shaped mouth-piece, the pitch being determined by the tension of the lips and by the resonance of the tube. The long coiled tube in such instruments has a very deep fundamental tone the numerous upper harmonics of which can be easily evoked. The cornet is also provided with little valves by which the effective length of the tube is varied and an additional number of tones made possible.

Vibration of Rods and Plates

355. Longitudinal Vibration of Rods. If a rod of steel, say a meter long and a centimeter in diameter, is held firmly at the middle point and if a cloth dusted with powdered rosin and folded over the rod is grasped firmly with the hand and drawn off the end with a quick strong pull, a clear, high-pitched sound may be produced due to the *longitudinal* vibrations of the rod. That the vibrations are of this nature may be demonstrated by means of a small ivory ball hung by a cord and resting against the end of the

rod. The ball will be violently driven off, swinging out as shown in the figure.

Glass tubes held at the middle may be similarly set in vibration, using a wet cloth instead of one dusted with rosin. Tyndall was able to set a large glass tube so powerfully in vibration by this means that the tube was shivered to pieces.

The middle point where the rod is held or clamped is a node and the ends vibrate lengthwise to and fro simultaneously toward the middle or away from it so that the bar is alternately lengthened and shortened. The vibrations are thus precisely like those in an open organ pipe where there is a node in the middle and a loop at both ends, and, as in the organ pipe, the period of a complete vibration is the time required for a compressional

FIG. 199. Ball driven from end of rod

wave to travel the length of the bar and back again. Thus if the velocity of sound or of compressional waves in steel is 5000 meters per second a bar 1 meter long will make 2500 vibrations per second, since the wave length is 2 meters.

The area of cross section of the bar does not affect the result, the same pitch is obtained from bars of different diameters and shapes of cross section if they are of the same material and length.

By a little dexterity such a rod may be made to give a higher harmonic, vibrating with a node in the middle, and two others, each one-sixth of the length of the bar from the end. The wave length in this case is evidently one-third of that in the former, and the frequency of vibration three times as great.

Rods of other metals or of wood or glass may be caused to vibrate in this way and the velocities of sound in them may be compared by their frequencies of vibration as shown by the tones which they give out.

356. Longitudinal Vibration of Wires. Longitudinal vibrations may also be set up in wires firmly clamped at both ends by rubbing them lengthwise with a bit of rosined cloth.

The clamped ends of the wires are nodes in this case and the middle is a loop. The pitch depends only on the velocity of sound along the wire and on its length and is quite independent of its tension except in so far as the tension affects the elasticity of the wire.

FIG. 200. Transverse vibration of rod

357. Transverse Vibrations of Bars. The *transverse* vibrations of bars are determined by their mass and stiffness, and hence depend on Young's modulus of elasticity, since it is this coefficient which determines a bar's resistance to bending. If a uniform free bar is struck at the middle point it tends to vibrate as shown in the figure, with a node near each end, and it may be supported at the nodes on wooden bridges without materially affecting its vibration.

In the xylophone or kaleidophone the bars of wood or metal vibrate transversely and are supported at their nodes.

358. Tuning-forks. A tuning-fork may be considered a bent bar vibrating in the mode shown in figure 200. For there are two nodal points, one on each leg of the fork near the bottom. The prongs swing alternately toward and away from each other, while the stem of the fork, being attached to the vibrating segment between the nodes, vibrates up and down. This is made apparent by the loud tone given out when the stem of a vibrating fork is touched to a wooden table top or sounding board.
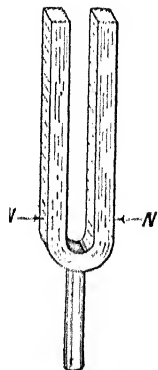
FIG. 201

Tuning-forks are often mounted on wooden resonators, boxes enclosing an air chamber capable of responding to the vibrations of the fork.

359. Law of Similar Systems. *When two vibrating systems are made of the same material and are exactly similar in dimensions, though not of the same size, their periods of vibration are proportional to their linear dimensions.* This law is shown by mathematical reasoning to be a consequence of mechanical principles, and is illustrated in many familiar instances.

For example, if two stopped organ pipes are constructed with cubical resonating chambers, but one having half the dimensions of the other, the smaller will vibrate with twice the frequency of the larger. Two tuning-forks of equally stiff steel and exactly similar in shape will be an octave apart if one is twice as large as the other. And so, also, if we take two straight steel bars, one of which has half the dimensions of the other in each direction, the smaller will make twice as many vibrations per second as the larger when vibrating in the same manner.

360. Vibration of Plates. The vibrations of flat plates of various shapes were studied by Chladni who scattered sand on the plates and observed the figures formed by the nodal lines in which the sand gathered when the plates were bowed. Some of
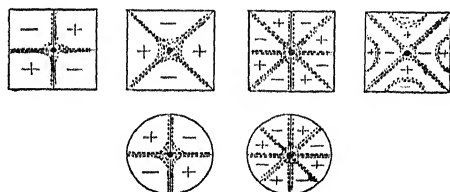
FIG. 202. Chladni's figures

these forms, known as Chladni's figures, are shown in figure 202. The upper row shows different modes of vibration that may be set up in a square plate supported at its center and bowed at some point on the edge. The slowest mode of vibration is the first, in which the vibrating segments are the four corners. *Segments separated by a nodal line must always be opposite in phase,* one vibrating up, while the other swings down. This opposition of phase is indicated by marking them alternately plus and minus.

If a resonator or wide-mouthed bottle which can respond to the vibrations of the plate is held with its mouth over any vibrating segment it will respond strongly, but if moved over a nodal line so that it is simultaneously acted on by two adjoining segments it is silent because the segments are in opposite phases.

So also when the plate is vibrating as shown in the first or second diagram in figure 202, if the hands are held just above two similarly vibrating segments so as to quench the sound waves coming off from them, the sound from the plate will be heard louder than before.

361. Bells. The blow of its tongue on a bell causes the circular rim to spring out into slightly elliptical shape, from which it springs back passing through the circular form into an ellipse with its greater axis at right angles to the first; thus it oscillates in four segments with four intermediate nodes as shown in the

figure. Making the rim of the bell thicker causes it to oscillate more quickly by reason of its increased stiffness and thus raises its pitch.

The above is its fundamental or slowest mode of vibration, but simultaneously with this the blow of the hammer sets up higher modes of vibration in which the rim may vibrate in 6, 8, or 10 segments with intermediate nodes. These higher tones are not
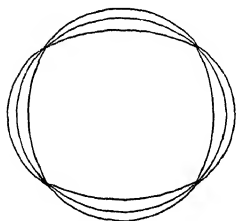
FIG. 203. Vibration of bell

in the harmonic series of the fundamental and hence the tones of bells are unsuitable for music. When the bell is first struck the higher tones are more prominent than the fundamental, but as the sound dies away the fundamental tone persists the longest.

The beating or throbbing heard as the tone of a bell dies away is due to want of uniformity in the rim, in consequence of which there are *two* fundamental tones of slightly different pitch. One or the other of these is excited according to the point struck by the hammer, though in general both are simultaneously set up.

MUSICAL RELATIONS OF PITCH

362. Musical Intervals Depend On Ratios. The musical effect of two tones when sounded together depends upon the *ratio of their frequencies*. This is well shown by means of the siren (§ 314). If four rows of holes in the siren are simultaneously used, in which the numbers of holes are proportional to 4, 5, 6, and 8, respectively, a combination of tones will be produced which will be recognized as the *major chord — do mi sol do*. And this musical relationship holds whatever may be the speed of the siren, showing that whatever the pitch may be it is the *ratio* of the frequencies of two tones which determines their musical relationship.

363. Harmonious Ratios. *Tones are harmonious whose frequencies are proportional to any two of the simple numbers 1, 2,*

3, 4, 5, 6. The most important harmonious ratios and their musical names are here given:

$$1 : 1 \text{ unison}$$
$$1 : 2 \text{ octave}$$
$$1 : 3 \text{ twelfth}$$
$$2 : 3 \text{ fifth}$$
$$3 : 4 \text{ fourth}$$
$$4 : 5 \text{ major third}$$
$$5 : 6 \text{ minor third}$$

The names are derived from the ordinary musical scale; thus the *octave* is the relation of the first and eighth tones of the scale; the *fifth*, that of the first and fifth; the *fourth*, that of the first and fourth, etc.

364. Major Scale. Three tones whose frequencies are in the ratio $4 : 5 : 6$ form what is known as a *major triad*.

The major scale is a sequence of tones so related that the first, third, and fifth tones form a major triad; also the fourth, sixth, and eighth, and the fifth, seventh, and ninth. The first note of the sequence is called the key-note and the triad starting with the key-note is the triad of the *tonic*. The fifth tone is known as the *dominant* and the fourth as the *subdominant*, and their triads are, respectively, known as the triads of the dominant and of the subdominant.

If the tones of the scale are represented by letters as in ordinary musical notation, their ratios for the key of C will be as follows:

DESIGNATION	C	D	E	F	G	A	B	c	d
Triad of tonic	4	..	5	..	6	..	..	..	..
Triad of subdominant	..	..	..	4	..	5	..	6	..
Triad of dominant ..	..	..	..	..	4	..	5	..	6
Ratio to tonic	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2	$\frac{9}{4}$

365. Tones and Half Tones. If the ratio of the vibration frequency of each tone to that of the one immediately preceding it is taken, we find

C		D		E		F		G		A		B		c		d
	$\frac{10}{9}$		$\frac{16}{15}$		$\frac{9}{8}$		$\frac{10}{9}$		$\frac{9}{8}$		$\frac{16}{15}$		$\frac{9}{8}$		etc.
	Tone		Tone		Half-tone		Tone		Tone		Tone		Half-tone		Tone

These ratios determine the musical character of the intervals. When the ratio of the frequencies of two tones is $\frac{9}{8}$ or $\frac{10}{9}$, they are said to differ a *whole tone*, while those whose ratio is $\frac{16}{15}$ are said to be a *half tone* apart.

366. Minor Triad. In the major triad, three tones whose ratios are 4 : 5 : 6, the interval between the first and second tone is a major third, while that between the second and third is a minor third. If we had three tones in the ratio 10 : 12 : 15, the interval between the first and second would be a minor third (5 : 6) while the interval between the second and third would be a major third (4 : 5). Such a combination of tones is known as a *minor triad*.

367. Minor Scale. A scale based on minor triads in the same way that the major scale is based on major triads is known as the minor scale. In the key of C the tones C, D, F, G are the same on both scales, while E, A, B each differs from the corresponding note of the major scale by the interval $\frac{25}{24}$, the minor tone being lower in each case. These tones of the minor scale may be designated E flat, A flat, B flat.

368. Temperament. Since there are two kinds of whole-tone intervals ($\frac{9}{8}$ and $\frac{10}{9}$) and also two kinds of half-tone intervals ($\frac{16}{15}$ and $\frac{25}{24}$), and since a note a half-tone higher than D, for example, which is called D sharp, would not be the same as E flat, it is clear that many notes are required to admit of playing music accurately even in a single key; and this number must be greatly increased if we are also to be able to play correctly in other keys.

In such an instrument as the violin the artist may indeed use true intervals, but in keyed instruments, like the piano or organ, the number of keys that would be required in such a case would make the key-board unmanageably complicated. Hence what is known as the *equally tempered* scale is used. *In this scale the whole tones are all equal and each equal to two half-tones.* And as there are five whole tones and two half-tones in an octave, the octave must be equivalent to six whole-tone intervals or twelve half-tone intervals; hence since two notes an octave apart are in the ratio 1 : 2, notes a whole tone apart must be in the ratio 1 : $\sqrt[6]{2}$, and those a half-tone apart in ratio 1 : $\sqrt[12]{2}$.

The following table gives in the upper row the vibration fre-

quencies of notes in the true or diatonic scale, beginning with middle C of the piano, while in the lower row are shown the corresponding frequencies in the equally tempered scale.

C	D	E	F	G	A	B	c	
261	293.6	326.2	348.0	391.5	435.0	489.4	522	Diatonic scale.
261	292.9	328.8	348.3	391.0	438.9	492.6	522	Equally tempered.

THE EAR AND HEARING

369. The Ear. To make clear the physical basis of hearing a short account of the structure of the ear will be required.

Referring to figure 204, three principal parts will be noticed: the *external ear channel* closed at the end by the tympanum or
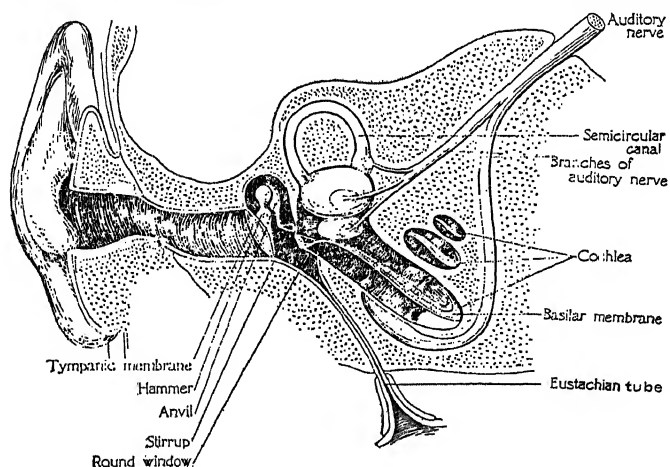
FIG. 204. Diagram showing tympanum, ossicles, and internal ear

drum skin, the *middle* ear in which is the chain of little bones or ossicles which connect the tympanum with the inner ear, and the *inner* ear itself in which the auditory nerve terminates and which is contained in a cavity in the massive part of the temporal bone. The small bones of the middle ear are situated in the upper part of a tube containing air, which is know as the *Eustachian tube* and which opens into the back of the mouth through a small valve

which opens in the act of swallowing. The air pressure in the Eustachian tube is thus kept the same as that on the outside of the drum skin. Persons going down in diving bells often experience a pain in the ears owing to the difference of pressure, which is relieved at once by swallowing.

The inner ear consists of a long chamber coiled up like a snail shell, and hence known as the *cochlea*, the *three semicircular canals*, and the *vestibule*. There are two openings from the Eustachian tube into the inner ear, one of which is closed by a membrane and the other by the stirrup bone or *stapes*, one of the four ossicles of the middle ear. The interior of the inner ear is filled with liquid, the *endolymph*.

The cochlea is probably that part of the ear by which musical sounds are distinguished in pitch and quality. It is divided into two parts by the *basilar membrane* which runs lengthwise through all its convolutions dividing it into two chambers. This membrane is strongly fibrous in structure, the fibers running across from one side of the tube of the cochlea to the other. Along its inner edge, where it is attached to the walls of the cochlea, the fibers of the auditory nerve terminate, so that a disturbance of any part of the basilar membrane causes a stimulus to the corresponding filament of the auditory nerve.

As this membrane also gradually varies in width from one end to the other, its fibers vary in length like the strings of a piano and have their own periods of vibration, which are slower for the long fibers and quicker for the shorter ones.

When sound waves falling on the tympanum cause it to vibrate, these vibrations are transmitted through the ossicles to the liquid on one side of the basilar membrane, then through the membrane itself to the liquid on the other side of it which is in contact with the flexible membrane closing the second opening into the Eustachian tube. The vibrations are transmitted most easily by that part of the basilar membrane which can vibrate in sympathy with the impressed vibration, and therefore the corresponding nerve filaments are stimulated. The arrangement is such that sounds of different pitch, awaking sympathetic vibrations in different portions of the basilar membrane, stimulate different nerve filaments and so give rise to different sensations.

It is now easy to understand why the ear should analyze com-

plex sounds, hearing each simple harmonic component as a separate simple tone. For it is in accordance with the mechanical laws of sympathetic resonance that when in a complex vibration there is a simple harmonic component which has the same period as the resonator, then the latter will respond. If there are, therefore, three different harmonic components in the vibrations communicated to the basilar membrane, the three corresponding portions of the membrane will be set in vibration, and consequently three different nerve filaments will be stimulated, exciting three distinct sensations of pitch.

370. Influence of Phase. From the above theory of audition developed by Helmholtz, it is to be expected that the relative phases of the components in a complex tone will have no influence on the resulting sensation, for the same parts of the basilar membrane are set in vibration whatever the phases of the component tones.

371. Beats. There is one important exception to this. *In case of beats the pulsations of tone are certainly due to the changing relative phases of the two components* which alternately act together and against each other. But this case is exceptional because *the two tones are so near in pitch that they affect closely adjacent portions of the basilar membrane.* It is not to be supposed that only a single fiber of the basilar membrane vibrates in response to a particular tone, but the adjoining portions are also set in vibration to some extent.

Suppose that two tones very near together in pitch are sounded and one excites the strongest response in the basilar membrane at *a*, figure 205, and the other at *b*.

The membrane on each side of *a* will also respond to the first tone and on each side of *b* to the second. If *a* and *b* are sufficiently near together, the fibers midway between them will vibrate simultaneously in sympathy
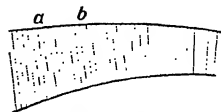
FIG. 205

with both tones; they will therefore take the resultant motion and will vibrate alternately strongly and feebly according as the two component vibrations are in the same or opposite phases. Hence the nerve filaments connected with these fibers receive an intermittent stimulus which produces the disagreeable jarring sensation of beats, just as the intermittent stimulus of a flickering light is painful to the eye (Helmholtz).

It is quite in accordance with this theory of audition that *rapid beats are heard as a distinct roughness and do not merge into a tone.* Thus if two Koenig forks, one making 2816 vibrations per second and the other 2560, are strongly sounded the beats are heard as an extremely disagreeable buzzing, though the number is 256 per second, while a *tone* of that frequency is wholly agreeable when sounded with either of the forks. The beating is due to the disturbance of the basilar membrane between the points where it responds to 2560 and 2816 vibrations per second while to excite the sensation of a tone having a frequency of 256 an entirely different portion of the membrane must be set in vibration, viz., that part which has a natural frequency of 256 per second.

372. Combinational Tones. Under some circumstances when two tones are strongly sounded, a tone is also heard whose frequency is equal to the difference between the frequencies of the two generating tones; its frequency is thus the same as that of the beats between the tones, though it is an entirely separate phenomenon. These *differential tones*, as they are called, may be very distinctly heard when two high-pitched forks are strongly sounded together, such as the two Koenig forks referred to in the last paragraph.

Helmholtz showed on mechanical principles that when two simple harmonic vibrations act on a membrane to set it in vibration, if the displacement of the membrane is so great that it is not simply proportional to the displacing force, then the resulting motion of the membrane will not be simply the sum of the two impressed harmonic vibrations, but will also include other components, one of which has a frequency equal to the *difference* of the frequencies of the two original vibrations, and the other a frequency equal to the *sum* of their frequencies.

Now, the tympanum of the ear is attached at the center to a small bone which is drawn inward by a muscle, thus keeping the drum skin tense, hence it is stretched in slightly conical form, and is therefore unsymmetrical and resists inward displacement more than it does outward. Hence, according to Helmholtz, when two strong vibrations are simultaneously impressed upon the tympanum, the motion which it communicates to the inner ear consists not simply of these, but includes also a vibration whose

frequency is their difference and another whose frequency is their sum. These are called the *differential* and *summational* tones. The latter were first observed by Helmholtz after he had shown theoretical reasons why they should exist. They are not so easily observed as the differential tones, and some observers have disputed their existence.

373. Helmholtz Theory of Dissonance and Consonance. Helmholtz showed that *dissonance was explained by beats taking place between either the tones themselves or their upper harmonics or the differential tones that they gave rise to,* and that *a clearly marked consonance occurs when the ratio of two tones is such that there are no beats, but when a slight change in the ratio gives rise to disagreeable beating.*

For example, the most perfect consonance is *unison,* for then the fundamental tones and upper harmonics all agree and there is no beating, but a slight mistuning causes beating not only between the fundamental tones, but between each pair of harmonics. Suppose, for example, one tone and its harmonics have the vibration frequencies shown in the series

<p style="text-align:center">100   200   300   400   500</p>

If the other tone, instead of making exactly 100 vibrations, makes 106, then it with its harmonics will form the series

<p style="text-align:center">106   212   318   424   530</p>

and there will be 6 beats per second between the fundamental tones, 12 between the first harmonics, 18 between the second harmonics, etc. The ear, therefore, selects a unison as a well-marked consonance. So also with the *octave:* suppose two tones which with their harmonics are given by the two series

	100	200	300	400	500	600
Octave		200		400		600

Here the fundamental of the tone making 200 vibrations per second is of exactly the same pitch as the first harmonic of the other tone, and there are no beats between any of the harmonics. But suppose the octave is mistuned, as, for example, below:

	100	200	300	400	500	600
Mistuned octave		210		420		630

Here there are 10 beats per second between the fundamental of one and the first harmonic of the other, 20 per second between the harmonics 400 and 420, etc., and the result is great dissonance.

It is clear from the above that the richer tones are in harmonics, the more dissonant they will be when mistuned. It is thus much easier to judge whether an octave is tuned correctly in case of two reed pipes than with two wide stopped pipes almost free from harmonics.

PROBLEMS

1. How long must a water wave be to travel with a velocity of 20 miles per hour?

2. What relation is there between the lengths of two water waves one of which has twice the velocity of the other?

3. Find the velocity of sound in dry air at 20° C. and pressure 73 cm. of mercury, when its velocity at 0° C. and 76 cm. pressure is 332 meters per sec.

4. What must be the amplitude of motion of the particles in a water wave, if the velocity of the particles at the wave crest is equal to the velocity of the wave?

5. If the height of a water wave from crest to trough is 3 ft. and its length is 50 ft., find its velocity, its frequency or the number of waves that pass per sec., and the direction and amount of the velocity of the water particles on the crest of the wave.

6. How many vibrations per second will be received from a bicycle whistle giving out 500 vibrations per sec. and approaching at the rate of 10 miles per hour?

7. If an observer were to move with the velocity of sound toward a sounding body at rest, what pitch would be heard? What if the observer were at rest while the sounding body approached him with the velocity of sound?

8. A tuning-fork having a frequency of vibration of 1000 per sec. is moved away from an observer and toward a flat wall with a velocity of 5 meters per sec. Find how many beats per second will be heard by the observer.

9. A cord 30 ft. long is stretched between two fixed supports with a force of 40 pounds' weight. How many transverse vibrations per sec. will the cord make if it weighs $\frac{1}{2}$ lb.?

10. A very long cord weighing 5 gms. per meter and stretched with a weight of 5 kgs. has one end made to oscillate sidewise 4 times per sec. Find the length of the waves set up in the cord.

11. A brass wire and a steel wire of the same diameter are stretched by equal weights and their lengths adjusted to give the same pitch when vibrating transversely. When the steel wire is 1 meter long between supports, how long will the brass wire be?

12. How many vibrations per sec. will be given out by an open organ pipe 76 cm. long? Give also the frequencies of its first three upper harmonics. Take temperature of air as 20° C.

13. How long must a stopped organ pipe be in order to have the same frequency of vibration as the open pipe in problem 12; also what are the frequencies of its first three upper harmonics?

14. What rise in temperature would raise the pitch of a flute pipe in an organ one semitone? Take original temperature as 0° C. and semitone ratio as $\frac{16}{15}$.

15. In a Kundt's tube filled with air the distance between the dust heaps is 17 cm., but when the tube is filled with carbon dioxide gas, the distance between nodes is 13.4 cm. Find the velocity of sound in the carbon dioxide if that in air is 340 meters per sec., both gases being at 15° C., and vibrations being produced by the same rod in both cases.

REFERENCES

TYNDALL: *Lectures on Sound* (Appleton). A delightfully written exposition of the subject by a brilliant experimenter and lecturer.

HELMHOLTZ: *Sensations of Tone*, translated by Ellis (Longmans). A thorough treatise on the scientific basis of music.

HEAT

374. Temperature Sense. The idea of temperature is obtained directly from our sense of touch. We speak of bodies as hot or cold according to the way in which they affect our temperature sense; and though temperatures cannot be accurately compared in this way, we may yet roughly estimate whether one body is hotter than another or whether a body is growing warmer or colder.

375. Transfer of Heat. When a hot body is brought into contact with a cold body the former is cooled while the latter is warmed. When a layer of copper is interposed between the hot and cold bodies the change goes on rapidly, but when a layer of felt is interposed the change is much slower. Hot water in a thermos bottle changes its temperature very slowly indeed, so that it is easy to imagine an ideal receptacle in which no change whatever in temperature could occur.

These facts indicate that the temperature of a body changes only when something passes into it from without or escapes from it to other bodies. This something is called heat.

Heat is said to pass from the hot to the colder body rather than that cold passes from the cold to the hot body, because experiment shows that when a body cools it loses something, namely, energy or power to do work, and hence heat rather than cold is considered the entity.

376. Other Effects of Heat. As bodies change in temperature other accompanying changes take place. As they grow hotter they increase in size, an enclosed mass of gas or vapor exerts a greater pressure, if heated enough a solid melts to a liquid, or a liquid is changed to vapor; also the elastic, electric, and magnetic properties of substances are seriously modified.

377. Equal Temperatures. When two bodies are placed in contact and no change takes place in either one such as would indicate a transfer of heat, they are said to be at the same tem-

perature. When one grows hotter and the other colder, the latter is said to be at a higher temperature than the other.

Temperature may be defined as that property of a body which determines the flow of heat. If there is no transfer of heat between two bodies when placed together, they are at the same temperature.

Thus temperature plays the same part in the flow of heat that pressure does in the flow of fluids.

378. Thermometers. To accurately compare temperatures instruments called *thermometers* are employed. Thermometers may be based on the expansive effect of heat, on the changes in pressure in a gas or vapor that are produced by change of temperature, on changes in the electrical properties of bodies, or, in short, on any easily measurable property of a substance, which depends on temperature.

Ordinary thermometers depend on the expansion of a liquid, such as mercury, alcohol, or ether, contained in a bulb of glass having a long tube or stem in which the liquid rises or sinks as it expands or contracts.

379. Fixed Points. In order that temperature observations by different observers may be comparable, *all thermometric scales are based on two fixed temperatures. These are the temperatures at which ice melts and that at which water boils under standard atmospheric pressure.*

FIG. 206. Boiling-point apparatus

1. The chief precaution to be taken in determining the freezing point is to see that the ice is free from salt. Ice from ponds frequently has traces of salts from the soil. Changes in barometric pressure do not affect the freezing point of water by as much as 0.001° C., and may, therefore, be disregarded.

2. The boiling point is determined by the use of some appa-

ratus, such as that shown in the figure, so that the whole thermometer up to the point at which the mercury stands in the stem is bathed in steam as it escapes from the boiling water. The escaping steam is made to pass down around the outside of the vessel so as to prevent the steam in contact with the thermometer from being cooled.

Impurities in the water may cause it to boil at a temperature slightly above the point at which pure water boils, but the escaping steam will have the same temperature as that from pure water if the pressure is the same. It is for this reason that the thermometer bulb is kept in the steam and is not allowed to dip into the water itself.

The boiling point is decidedly influenced by changes in atmospheric pressure. *An increase of* 27.0 *mm. in the barometric height raises the boiling point by one whole degree Centigrade.*
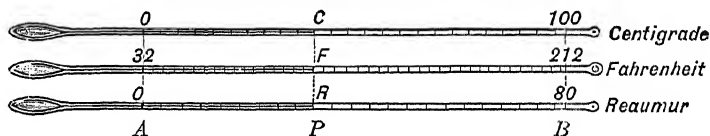
FIG. 207.   Thermometric scales

380. Scales of Temperature. In order that a thermometer may be useful in determining intermediate temperatures it must be graduated or divided into intervals or degrees.

Three scales are in general use: the Centigrade or Celsius scale, used extensively on the Continent and in most scientific investigations; the Fahrenheit scale, used chiefly in English-speaking countries; and that of Reaumur, used to some extent on the continent of Europe.

In the Centigrade scale the freezing point is marked zero and the boiling point 100, the interval being divided into 100 degrees.

In Fahrenheit's scale the freezing point is 32° and the boiling point is 212°, so that there are 180 degrees between the two.

In Reaumur's scale the freezing point is 0° and the boiling point 80°. The relation of the three scales is shown in figure 207.

Since 180 Fahrenheit degrees correspond to 100 Centigrade degrees, a Fahrenheit degree is $\frac{5}{9}$ of a Centigrade degree. To change Fahrenheit temperatures to Centigrade we, therefore, subtract 32° and take $\frac{5}{9}$ of the

remainder. On the other hand, to change Centigrade temperatures to the Fahrenheit scale add 32° to $\frac{9}{5}$ of the Centigrade temperature.

Or since the ratio of the number of degrees between A and P (Fig. 207) to the whole number of degrees between A and B is the same in each case, we have the equations

$$\frac{C}{100} = \frac{F - 32}{180} = \frac{R}{80}$$

by which relations temperatures on any one scale may be changed to either of the others.

381. Graduation of the Scale. A thermometer is ordinarily graduated so that each degree corresponds to an equal apparent increase in the *volume* of the mercury. Thus if the thermometer tube is perfectly cylindrical the degree marks are equidistant, but if the tube is not uniform in diameter the degree marks should be so spaced that the *volume* between consecutive marks is the same everywhere throughout its length. The graduations will, therefore, be closer together where the tube is wider and farther apart in the narrower portions of the tube. The 50° mark should be so placed that it divides in half the volume between the 0° and 100° points.

382. Arbitrary Feature of Thermometric Scales. But it is evident that even such a scale depends upon the properties of the expanding substance and if we were to take two thermometers, one containing alcohol and the other mercury, and were to graduate them in this manner, while they would agree at the fixed points they might not agree anywhere else.

This arbitrary element enters into every scale of temperature. Suppose, for example, we were to define 50° as that temperature which results from mixing equal weights of water at 0° and 100°, respectively. It would be found that if mercury had been taken instead of water the resulting temperature would have been different. And even if we were to define 1° as the rise in temperature of a given mass of water due to the addition of $\frac{1}{100}$ of the whole amount of heat required to raise it from 0° to 100°, the scale of temperature obtained would be different from that obtained by using in a similar way some other substance than water.

383. Peculiarities and Defects of Thermometers. The rise of the mercury in a thermometer when it is heated is due to the *difference* between the expansion of the mercury and that of

the glass bulb; for if the mercury and glass expanded equally the mercury would not rise at all in the tube. Therefore, two mercurial thermometers may not agree except at the fixed points, unless they are made of the same kind of glass.

The most serious defect in mercurial thermometers is the change in the zero point. When the bulb is cooled after having

FIG. 208. Six's maximum and minimum thermometer

been heated it takes a long time to return to its original dimensions. Thus if a thermometer is heated to 100° and then quickly cooled, the zero point will be found lower than before it was heated; this is known as the *depression* of the zero point. The bulb continues slowly to contract, but it may be weeks before the original zero point is reached.

This depression of the zero point may amount to two- or three-tenths of a degree and is a source of error that affects more or less all temperature observations with such instruments, for the reading at a particular temperature depends on whether or not the thermometer has recently been heated to a higher temperature.

Researches carried on at Jena have resulted in the production of a special glass for thermometers, known as the Jena normal glass, which is almost free from this defect and is, therefore, employed in making thermometers for exact work.

384. Spirit Thermometers. Alcohol and ether thermometers can be used at temperatures so low that mercury would freeze, their expansions also are so much greater than mercury that so fine a tube is not required. But these liquids wet the tube and if the upper part of the stem is cooler than the surface of the liquid column the liquid will distil and condense in the upper end of the tube.

Alcohol expands 6 times as much as mercury and ether $8\frac{1}{2}$ times as much, they are therefore suitable for sensitive thermometers, though on account of the pressure of the vapors of these liquids an alcohol thermometer should not be used above 100° C. and an ether thermometer not above 60° C.

385. Maximum and Minimum Thermometer. For registering maximum and minimum temperatures the form of instrument devised by Six (Fig. 208) is found convenient. In this instrument the bulb B and the

tube as far as the mercury column at C is filled with phenol or some liquid having a large expansion coefficient. The mercury column fills the lower part of the tube between C and D while the tube above D is also filled with phenol reaching up to the bulb A which is only partly filled. When the temperature rises the expansion of the liquid in the bulb B causes the mercury column to sink at C and rise at D, pushing upward a little index of iron in the tube above D which in consequence of friction remains where it is pushed and marks the maximum temperature. On cooling the contraction of the liquid in B causes the mercury to rise at C pushing upward a little index at that point which marks the minimum temperature. To set the instrument the indices are drawn down against the mercury column by means of a small magnet.

386. **The clinical thermometer** used by physicians is a maximum thermometer having a short scale ranging from about 95° to 108° F.; the tube is made very flat and narrow just above the bulb. The mercury will readily pass through the constriction in rising, but as it contracts capillary force causes the column to separate at that point, leaving the upper part of the mercury column to mark the maximum point. To set the instrument the mercury is brought back to the bulb by a vigorous shake.

387. **Air Thermometer.** Since mercury-in-glass thermometers can be used only between −40° C. and 450° C., and their readings are so much influenced by the peculiarities of the kinds of glass of which they are made, they are not suited to be used as independent standards. For standard purposes air or hydrogen or nitrogen may be used as the thermometric substance, because with a porcelain bulb such a thermometer can be used from − 200° C. up to 1500° C., and the expansion of these gases is so great (more than twenty times that of mercury)

Fig. 209. Galileo's air thermometer

that the expansion of the glass or porcelain bulb containing the gas is quite insignificant in comparison, and can be allowed for without sensible error.

A crude form of air thermometer which is interesting because it was used in 1597 by Galileo, the inventor of the thermometer, is shown in figure 209. The bulb containing air terminates in a tube dipping into a vessel of colored liquid which rises or sinks in the tube, according as the enclosed air contracts or expands. The most evident defect of this arrangement is that the liquid column will rise and fall as the atmospheric pressure

changes, even though the temperature of the gas may remain constant.

388. Standard Air Thermometer of Constant Volume. For the exact measurement of temperature by the air thermometer it is found most convenient to keep the *volume* of the air constant and use its *pressure* to measure temperature. A form of instrument devised by Jolly is much used. The bulb A contains the gas to be used, which may be hydrogen or nitrogen or air that has been dried and freed from carbon dioxide. The bulb is connected by a capillary tube with the wider tube at B.

The vertical tubes B and DE are connected by a flexible rubber tube, which is full of mercury, the mercury column extending up into the glass tubes at B and E. The tube DE is attached to a slide and can be raised or lowered along a fixed scale and clamped at any point.

In using the instrument the air in the bulb is first cooled to $0°$ in melting ice and the tube DE adjusted in height until the mercury at B comes exactly to a fixed mark at the end of the capillary tube. The pressure of the enclosed air is then obtained by subtracting the height of the mercury column BE' from the barometric height which gives the pressure of the external air on E. In a similar way the pressure of the enclosed air may be measured when the bulb is heated to $100°$ in steam. In this case also the mercury level at B must be adjusted to the same point as before, keeping the volume of the air constant except for small changes in the size of the bulb itself.

FIG. 210. Jolly's air thermometer

The pressures of the enclosed gas in these two cases may be represented by p_0 and p_1, respectively.

If it is now desired to determine the temperature of a bath in which the bulb is immersed it is only necessary to measure the pressure p exerted by the gas just as in the other cases. If this pressure is found to be half-way between p_0 and p_1 the temperature of the bath is $50°$. Or, in general, if t is the temperature to be determined corresponding to the pressure p

$$t : 100 :: p - p_0 : p_1 - p_0.$$

In the most refined work we must make corrections for the expansion of the glass bulb itself due both to changes in temperature and pressure, and also take account of the fact that the gas just above B is not at the same temperature as the bulb A.

Such a process would evidently be too cumbrous to employ except for the purpose of standardizing some more convenient working form of instrument, such as the mercurial thermometer or the electrical resistance thermometer.

389. Electrical Methods. Some very important methods of measuring temperatures are based on electrical phenomena and will be more particularly described in that connection.

The thermoelectric method determines temperature by measuring the electromotive force set up when the junction of two wires made of different metals is heated. For low temperatures a copper-iron junction may be used, while for high temperatures the junction of a pure platinum wire with one of the platinum-rhodium alloy is used.

The resistance method depends on the increase in the electrical resistance of a coil of pure platinum wire with rise in temperature (§ 663).

A resistance thermometer consists of a coil of platinum wire mounted in a glass or porcelain tube to protect it from injury and contamination, and provided with connections by which its electrical resistance may be tested. By means of suitable accessory apparatus the temperature of the coil may be read directly without calculations. On account of the range of temperatures that can be measured in this way (from $-270°$ to $1500°$ C.), and the accuracy and ease with which the determinations may be made, this is one of the most valuable of all methods of temperature measurement.

390. High Temperatures. For measuring high temperatures the gas thermometer, or electrical methods, or radiation pyrometers may be used.

The hydrogen gas thermometer having a porcelain bulb may be used up to $1500°$ C. (§ 388).

The platinum-rhodium thermo-couple and the electrical resistance thermometer may be used up to $1500°$ C. if protected by porcelain tubes.

For the highest temperatures *radiation pyrometers* are used. These are of two types. One depends on the heating power of the radiation from a mass of molten metal or from the interior of a furnace, and is so devised that it may be used at quite a distance from the hot body if the radiating surface is large.

The other type depends on a measurement of the intensity of the light from the glowing hot body or interior of a furnace.

PROBLEMS

1. Find the Fahrenheit temperatures corresponding to $80°$, $20°$, $-10°$, and $-50°$ C.

2. Find the Centigrade temperatures corresponding to $1000°$, $98.6°$, $0°$, and $-50°$ F.

3. What temperature reads the same on both Fahrenheit and Centigrade scales, and at what temperature is the Fahrenheit scale-reading twice that on the Centigrade scale?

4. A temperature interval of 35° on the Centigrade scale is an interval of how many degrees Fahrenheit?

5. The absolute zero of temperature is $-273°$ on the Centigrade scale; what is it on the Fahrenheit scale?

6. Calculate the Fahrenheit temperatures of the melting points of iron, copper, lead, and mercury. (See p. 298, in §434.)

Expansion of Solids

391. Expansion of Solids. Almost all solids expand when heated. Isotropic bodies, such as glass and all liquids, expand equally in every direction. Crystals in general expand differently in different directions, and may even contract along one direction and expand in another, but in most cases the expansion more than makes up for the contraction so that there is on the whole an increase in volume with rising temperature.

392. Coefficient of Linear Expansion. *The fractional part of its length that a rod elongates when raised one degree in temperature is called its coefficient of linear expansion.* Let the length of a bar at 0° be l_0, and let a be its coefficient of linear expansion, then its increase in length for a rise in temperature of 1° will be l_0a, and for t degrees its increase in length is l_0at, so that its total length l at the higher temperature is:

$$l = l_0 + l_0at \quad \text{or} \quad l = l_0(1 + at).$$

In this formula l may be taken as the length of the bar at a temperature t degrees higher than that at which its length is l_0, even though the latter may not be its length at 0° C.

It must not be supposed that the coefficient of expansion of a substance is the same at all temperatures, for in general it increases as the temperature rises. In the above formula a represents the average value of the coefficient throughout the rise in temperature represented by t.

393. Coefficient of Volume Expansion. If a cube of substance is taken measuring 1 cm. each way at 0°, and having a coefficient

of linear expansion a, then its linear dimensions at $t°$ will be $1 + at$ and its volume will be

$$(1 + at)^3 = 1 + 3at + 3a^2t^2 + a^3t^3,$$

but the coefficient a is so small that the terms involving a^2 and a^3 may be neglected and the volume may be expressed as

$$1 + 3at.$$

$3a$, therefore, represents the increase in volume of a unit cube for one degree rise in temperature and may be called the coefficient of cubical or volume expansion; hence *the coefficient of cubical expansion is three times the coefficient of linear expansion in an isotropic body*.

394. Measurement of Coefficients of Expansion. When the substance whose coefficient of expansion is to be obtained
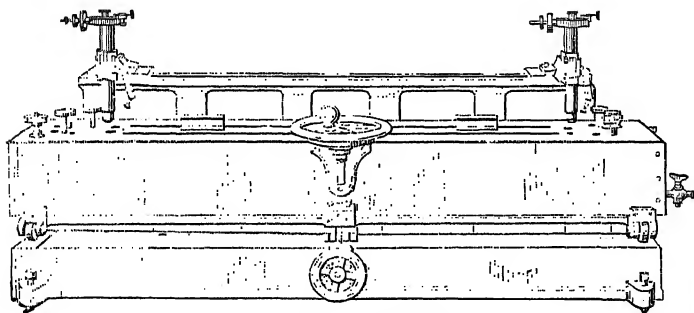
FIG. 211. Comparator

has the form of a long rod, its expansion may be measured by a comparator such as that shown in the figure.

Two microscopes are set on two marks on the bar, one near each end. The microscopes are firmly clamped to a solid base which is kept free from temperature change. The bar to be examined is enclosed in a box provided with glass windows through which the microscopes are set on the marks. The bar is first packed in melting ice and the micrometers attached to the microscopes are set on the two marks. Then water at a higher temperature is caused to circulate through the box, maintaining a constant higher temperature, and the micrometers are again set on the two marks. The difference between the

micrometer readings gives the elongation of the bar and accurate thermometers give the change in temperature. The whole length of the bar between the marks is then carefully determined.

If this length is l and the elongation is e when the temperature is raised from t to t', the coefficient of expansion a is found from the relation $e = la(t' - t)$, or

$$ a = \frac{}{l(t' - t)}. $$

This is the average value of the coefficient between the temperatures t and t'.

395. Expansion of Crystals. Crystals that do not belong to the *regular system* expand differently in different directions. A sphere cut out of such a crystal will become an ellipsoid when its temperature is raised. In some cases two of the axes of the ellipsoid would be found of the same length and in some cases all three would be different. The directions in the crystal corresponding to the axes of the ellipsoid are called the axes of thermal expansion. In quartz the expansion at right angles to the axis of the crystal is nearly twice the expansion in the direction of the axis.

TABLE OF COEFFICIENTS OF LINEAR EXPANSION, PER DEGREE CENTIGRADE

Invar.......	.00000096	Brass........	.0000189
Glass.......	089	Silver........	194
Platinum....	089	Aluminum....	222
Steel........	110	Lead.........	280
Iron........	117	Zinc........	298
Copper.....	167	Ebonite......	770

These values are approximate. The exact value for any substance depends on the state of hardness, purity, and temperature of the specimen.

396. Some Illustrations. An iron tire when heated expands so that it can easily be slipped over the wooden rim of the wheel, which it binds firmly on cooling. So the breeches of cannon are strengthened by having a series of tubes shrunk over the inner core, in this way producing an outside compression of the core which enables it to withstand the enormous pressure of the powder gas.

Allowance has to be made for expansion in case of bridges. In a steel bridge 1000 ft. long the change in length between extremes of summer and winter may amount to 8 in.

The aggregate length of the rails in a mile of track may be 4 ft. longer when hottest than when coldest, so that an allowance of about 0.3 of an inch is needed for each 30-ft. rail. The grate bars of furnaces rest loosely in their supports in order to allow expansion, and long steam pipes are provided with sliding or " expansion " joints unless the bends in the pipe are such as to yield elastically to elongation and contraction.

Quartz crystals have very large expansion, and when unequally heated fly to pieces because of the great strains which result in that case. When quartz is fused, however, into a glass, its coefficient of expansion is extremely small, and vessels made of fused quartz may, when red hot, be suddenly quenched in water without breaking.

A specially prepared nickel-steel, having 36.1 per cent of nickel and known as *invar*, has a temperature coefficient of only 0.0000009 or $\frac{1}{10}$ as large as platinum. It is of great value for measuring bars and tapes, and for pendulums.

Whenever wires are to be hermetically sealed into glass, as in the case of the connections of an incandescent lamp filament, the wire must have a coefficient of expansion very nearly equal to that of glass. Wires of a metal having a greater coefficient of expansion would shrink away from the glass on cooling, leaving a crack through which air could pass. Platinum was at one time used for this purpose as this metal was the only one known with a coefficient of expansion equal to that of glass, but an alloy of iron and nickel called *platinum substitute* is now used in place of platinum.

397. Compensated Clock Pendulums. The elongation of a clock pendulum with rising temperature causes it to swing more slowly and the clock loses time. Dry wood pendulum rods have very small expansion and so are sometimes used, but they are affected by moisture. For the most accurate clocks compensated pendulums are used. One of the best forms is *Graham's mercurial pendulum* (Fig. 213), where a reservoir of glass or steel containing mercury is hung by a steel rod. If properly designed, the raising of the center of oscillation (§ 148) due to expansion of the mercury is balanced by the lowering due to the elongation of the steel suspending rod, so that the effective length remains constant.

In Harrison's *gridiron pendulum* (Fig. 212) the expansion of the steel bars *FF* will lower the bob, while the expansion of the brass rods *CC* will tend to raise it. If the upward elongations of *C* and *C* for a given change

in temperature are together equal to the combined downward elongations of *FFF* the bob will neither be raised nor lowered.

398. Watch Compensation. The balance-wheel of a watch if uncompensated will run slower as the temperature rises, because the elasticity of the hair-spring is less at higher temperatures, and also the expansion of the wheel makes its moment of inertia greater.

Compensation is secured by making the balance-wheel as shown in figure 214. The rim is made of brass on the outside and steel on the inside, and instead of being continuous it is cut in two segments which are connected rigidly by a cross-
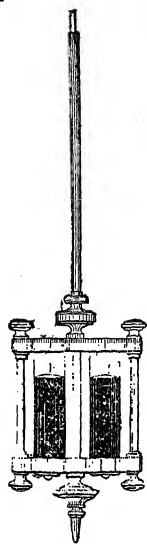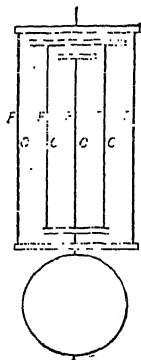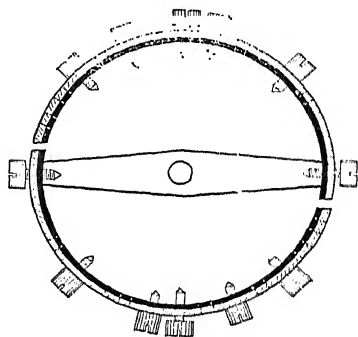
FIG. 213   FIG. 212   FIG. 214

bar. When the temperature rises the brass outer side of the rim expands more than the steel inner side so that the free ends of the segments bend inward, thus carrying part of the mass in toward the axis and so tending to compensate the outward expansion of the cross-bar, and the diminished elasticity of the hair-spring. The adjustment is completed by means of little screws set in the rim of the wheel. Those near the free points tend to increase the compensation, while those near the fixed ends of the segments have the opposite effect.

399. Force of Contraction. The force produced by the shrinking of a bar on cooling is the same as would be required to stretch it by the same amount at the same temperature.

PROBLEMS

1. What is the change in length of the steel cables of a suspension bridge 2000 ft. long between the extremes −20° F. and 97° F.?

2. A brass meter bar is correct at 15° C.; what will be its length at 20° C.?

3. What is the coefficient of expansion of a 30-ft. steel rail on the Centigrade scale and also on the Fahrenheit scale if it changes in length 0.234 in. when the temperature ranges from $-17°$ F. to $100°$ F.?

4. At $20°$ C. a brass plug 5 cms. in diameter is $\frac{1}{100}$ of a millimeter too large to fit a hole in a steel plate. At what temperature will it just fit?

5. A glass specific-gravity bottle has a capacity of exactly 300 c.c. at $15°$ C.; what will be its capacity at $0°$ C.?

6. A cylindrical zinc pendulum bob has a hole running lengthwise through it in the direction of its axis through which the steel pendulum rod passes, and rests on a cross-piece at the lower end of the rod. How long must the rod and the bob be that the center of gravity of the bob may remain constant at 95 cms. below the point of support while the temperature changes? Take expansion coefficient of steel as 0.000010 and for zinc 0.000029.

EXPANSION OF LIQUIDS

400. Expansion of Liquids. When a liquid contained in a bulb provided with a long neck is heated, it rises in the stem by *an amount which depends on the difference between the expansion of the liquid and that of the bulb.* The rise indicates what is known as the *apparent expansion.* If a bulb containing liquid is suddenly plunged into a vessel of hot water the liquid in the stem may be observed to sink at first because the bulb expands before the liquid within is fully heated.

To determine the expansion of a liquid take a bulb with a graduated stem like the tube of a thermometer and calibrate the stem, or determine the relation between the volume of the whole bulb and the volume of the divisions of the stem. This may be done by filling the bulb with mercury and weighing it, and then separately weighing the amount of mercury required to fill a certain number of divisions of the stem; the relative weights give the relation between the volumes. The bulb is now filled with some liquid up to a certain mark on the stem and then packed in

FIG. 215. Bulb with graduated stem

ice or cooled to some steady low temperature and the point to which the liquid contracts is observed. It is then warmed to some higher temperature and the point at which the liquid stands

is again observed. From the divisions of the stem between these two points the apparent increase in volume is determined, and if this is divided by the original volume and then by the rise in temperature, the *apparent coefficient of expansion* is obtained.

The expansion of a glass bulb of volume V, is Vat where a is the coefficient of volume expansion of glass and t is its rise in temperature, while the expansion of the contained liquid is Vbt where b is its coefficient of expansion.

Since the rise of the liquid in the stem is due to the excess of its expansion over that of the bulb the apparent expansion is

$$Vbt - Vat = Vt(b - a).$$

The apparent coefficient of expansion is therefore $b - a$, or the difference between the coefficients of expansion of the liquid and the bulb.

Hence the coefficient of expansion of the bulb must be determined before that of the contained liquid becomes known. This
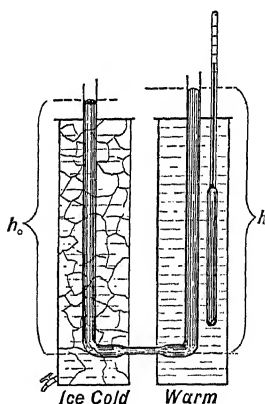
FIG. 216.   Expansion of mercury

may be accomplished either by studying the expansion of a bar made of identically the same glass or by observing the apparent expansion in the bulb of some liquid whose coefficient of expansion is already known.

401. Absolute Expansion of Mercury. The expansion of mercury has been studied with great care because it is the liquid best adapted for use in determining the coefficients of expansion of bulbs to be used in the study of other liquids. Its coefficient of expansion was determined by Dulong and Petit by the following method which is *independent of the expansion of the tube containing the mercury.*

Two vertical tubes (Fig. 216) connected at the bottom by a very thin horizontal cross tube, contain mercury. One is packed in ice and the other is heated to some known temperature t. Then by the laws of hydrostatics the less dense liquid will stand higher, and the height of the cold column of mercury

multiplied by its density is equal to the product of the height of the hot column by its density, or

$$hd = h_0 d_0. \tag{1}$$

But as a given mass of mercury expands in volume it diminishes in density, so that

$$V : V_0 = d_0 : d, \tag{2}$$

and since

$$V : V_0 = 1 + at : 1$$
$$d_0 : d = 1 + at : 1$$

or

$$d_0 = d(1 + at)$$

and by equation (1)

$$\frac{h}{h_0} = 1 + at. \tag{3}$$

So that by measuring the heights h and h_0 and determining the temperature t of the hot column the coefficient of volume expansion a of the mercury becomes known.

402. Expansion of Water. The expansion of water has been determined with great accuracy at the German National Laboratory or *Reichsanstalt* by the method just described.

The curve of expansion (Fig. 217) shows that when water is
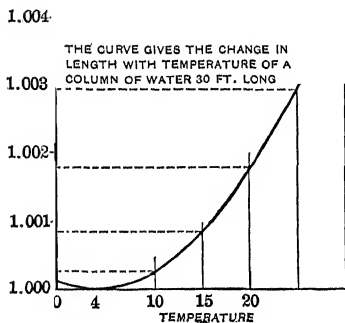
FIG. 217. Expansion of water from 0°–30°
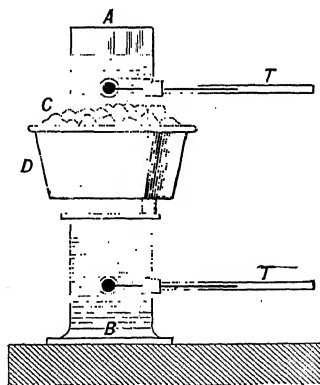
FIG. 218. Hope's apparatus for determining the temperature of maximum density of water

heated from 0°, it first contracts and then expands, reaching its maximum density at almost exactly 4° C.

This fact is of great importance in nature, for the cooling of a lake goes on rapidly at first, the cooled surface water settling to the bottom, thus aiding the cooling of the whole by convection currents. But when the water has reached 4° C. any further cooling must be accomplished by the slow process of conduction, for the colder water being less dense will remain at the top. So ice forms at the top and only gradually thickens downward, and if the lake or pond is not too shallow the bottom does not fall below 4° C. for there is a small supply of heat flowing out from the earth which makes up for that lost by conduction toward the surface.

Hope made use of the apparatus shown in figure 218 to determine the temperature of maximum density of water. A vessel of water provided with thermometers at the top and bottom is cooled above by being surrounded by ice. The lower part of the vessel is carefully jacketed with cotton or felt to prevent the inflow of heat through the sides. The upper thermometer will at first stand higher than the other, but finally the lower will stand steadily at 4° C. while the upper will cool below that point.

The temperature of maximum density of water is lowered when salt is dissolved in it. Sea water attains its maximum density only at −3.67°, which is below its normal freezing point.

DENSITY AND VOLUME PER GRAM OF WATER

TEMPERATURE	DENSITY	VOLUME PER GRAM	
0° C.	0.999867	1.000132 c.c.	As found at the Reichs-anstalt by the method of balancing columns.
3.98	1.000000	1.000000	
10.00	0.999727	1.000272	
15.00	0.999126	1.000874	
20.00	0.998229	1.001773	
25.00	0.997071	1.002937	
30.00	0.995673	1.004345	
35.00	0.994057	1.005977	
40.00	0.992241	1.007819	
60.00	0.9834	1.0169	Approximate values.
80.00	0.9719	1.0289	
100.00	0.9586	1.0431	

COEFFICIENTS OF EXPANSION OF SOME LIQUIDS

	At 0°	At 20°	At 40°	Average Between 0–40°
Water	− 0.000067	0.000206	0.000388	0.000192
Mercury	0.000179	0.000180	0.000181	0.000180
Alcohol	....	....	....	0.00112
Ether	0.00151	0.00165	0.00189	0.00167

It will be observed that these coefficients are larger than those of solids, and that in general they increase with the temperature.

Expansion of Gases

403. Expansion of Gases. The expansion of gases with heat is much greater than that of solids or liquids and is remarkable for being *nearly the same for all gases*. On account of the great compressibility of gases there are two distinct conditions under which their expansion by heat may be determined. First, the pressure may be kept constant and the volume expansion of the gas measured as the temperature rises, or, second, the volume of the gas may be kept constant and the increase in pressure with rising temperature may be measured.

If the gas perfectly obeyed *Boyle's law* its coefficient of expansion at constant pressure would be equal to that with constant volume.

404. Expansion at Constant Pressure. Gay-Lussac was the first to

Fig. 219

carefully study the expansion of gases at constant pressure, but Regnault by the apparatus indicated in the diagram obtained far more accurate results.

The bulb *A* is filled with the gas to be studied and cooled to zero by means of melting ice. By the stopcock *E* it is then shut off from the gas supply and connected with *B* which is completely filled with mercury up to the opening of the small tube
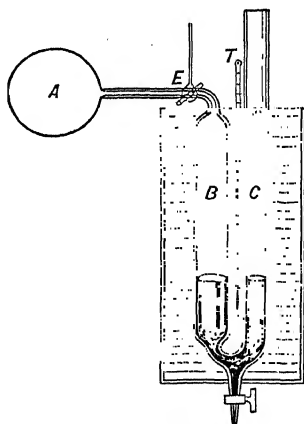
at the top, and if the gas in the bulb is at the same pressure as the outer air the mercury will stand in the open tube C at the same level as in B. The bulb A is then heated to any desired temperature, say to 100°, and as the gas expands mercury is allowed to flow out of the stopcock at the bottom so that it is kept at the same level in B and C, thus maintaining the pressure constant. Part of the expanded air is in A at 100° and part in B at the temperature of the water bath which surrounds the tubes. The tube B is graduated, so that the exact volume of the expanded gas may be determined.

405. Increase in Pressure at Constant Volume. Regnault also was the first to make accurate measurements of the increase in pressure of a gas when the volume is kept constant. The apparatus used is the same as that described above (Fig. 219) but when the bulb A is heated and the expanding gas begins to force down the mercury in B, more mercury is poured into C until the additional pressure again causes the mercury to exactly fill B. In this way the heated gas is kept confined in the bulb A, and its pressure is measured by the height of the mercury in C above that in B together with the height of the barometer. In this experiment the bulb is expanded slightly both by the rise in temperature and by the increased pressure in the interior, and on account of this change in volume a small correction must be applied.

COEFFICIENTS OF EXPANSION OF GASES

GAS	INCREASE IN VOLUME AT CONSTANT PRESSURE PER DEGREE C.	INCREASE IN PRESSURE AT CONSTANT VOLUME PER DEGREE C.
Air.....................	0.003671	0.003668
Oxygen.................	....	3674
Nitrogen...............	3671	3668
Hydrogen..............	3661	3660
Carbon monoxide........	3669	3667
Carbon dioxide..........	3710	3687
Sulphurous acid.........	3903	3845

One cubic foot of air at 0° would expand to 1.367 cu. ft. at 100°, an increase of more than $\frac{1}{3}$ of its volume at 0° C.

From the above table it is clear that *different gases have nearly*

equal coefficients of expansion. This is known as *the law of Charles or Gay-Lussac.*

The increase in volume of a gas per degree rise in temperature is about $\frac{1}{273}$ of its volume at 0° C.

406. Absolute Scale of the Air Thermometer. According to Charles' law, gases, at constant pressure, expand nearly 0.00366, or $\frac{1}{273}$ of their volume at zero for a rise in temperature of one degree Centigrade. Consider a cylinder filled with air or hydrogen and closed by a piston which always exerts the same pressure on the enclosed gas. When the gas is at 0° suppose the piston stands at A, then when the gas is warmed to 100° it expands and the piston rises to B. If we divide the space from A to B into 100 equal parts and continue the graduation down below A, marking off equal spaces for every degree, we shall find that there will be 273 degrees below the zero. If we now call the bottom of the cylinder the zero point we shall have a scale of temperature in which 273° will be the freezing point of water and 373° will be the boiling point. This scale is called the absolute scale of the air thermometer, and its zero is called the absolute zero. It is only necessary to add 273° to any Centigrade temperature to obtain the corresponding temperature on the absolute scale. It will be seen from the way in which the scale is obtained above, that the volume of the gas in the cylinder is proportional to its temperature on the absolute scale, and since all gases have nearly the same

FIG. 220

coefficient of expansion, it may be stated as true in general, for all gases that are not too near their points of condensation, that *the volume of a gas is very nearly proportional to its absolute temperature when the pressure is kept constant. So also when the volume is kept constant the pressure of a gas is nearly proportional to the absolute temperature.* At the absolute zero the pressure would be zero. There are good reasons for believing that the pressure of a gas is proportional to the energy of vibration of the molecules and therefore that at the absolute zero the molecules of gas have no energy of motion. Consequently this is the lowest
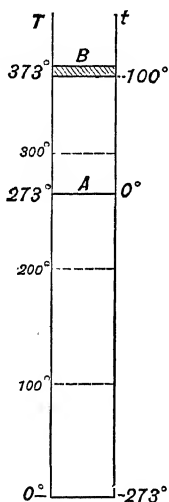
possible temperature, for if a substance has no energy of motion to give up, it cannot give out any heat and be cooled further.

Of course no gas would actually be reduced to zero volume, however much it might be cooled, though its pressure might be reduced to zero. It would condense into a liquid and cease to behave as a gas before reaching zero volume.

An entirely independent and more conclusive line of reasoning has led to the establishment of *the absolute thermodynamic scale* of temperature. (See Appendix I.) This is independent of the properties of any particular substance, and its zero is the lowest possible temperature. Experiment shows that the absolute scale based on the expansion of gases agrees almost exactly with the thermodynamic scale except at the very lowest temperatures. The zero of the gas scale is therefore properly called the *absolute* zero.

By means of liquid air, temperatures as low as $-200°$ C. may be obtained, and by the evaporation of liquid hydrogen $-258°$ C. has been reached, only $15°$ above the absolute zero. By the evaporation of liquid helium at low pressure the remarkable temperature of only $0.8°$ above the absolute zero has been reached by the Dutch physicist, Onnes. At these low temperatures rubber and steel become as brittle as glass, lead becomes stiff and elastic, while the electrical resistances of metals are greatly reduced.

407. General Gas Formula. As was shown in the last paragraph, the volume of a given mass of gas kept at constant pressure is proportional to its temperature on the absolute scale; that is,

$$\frac{v}{T} \qquad \frac{v_0}{T_0}$$

where $T = 273 + t$ and $T_0 = 273 + t_0$, T and T_0 being the absolute temperatures corresponding to t and t_0 of the ordinary Centigrade scale.

By taking account of Boyle's law also, it may be shown that, *in general, whatever changes may take place in the pressure volume and temperature of a given mass of gas, the initial pressure volume and temperature are connected with their final values by the relation*

$$\frac{p_0 v_0}{T_0} = \frac{pv}{T}.$$

For, suppose a given mass of gas in a cylinder is in the state A having volume v_0, pressure p_0 and temperature T_0, and is to be brought into a state C in which the volume, pressure and temperature are all changed. Let the gas first have its temperature raised to T, *keeping the pressure constant at p_0*, it will come to a volume v' such that

$$\frac{v'}{T} = \frac{v_0}{T_0}$$

by Charles' law. Now *keep the temperature constant* at T and change the pressure to p, the volume will change from v' to v, and by Boyle's law we have $pv = p_0 v'$ or

$$v' = \frac{pv}{p_0}$$

Substituting this value in the previous equation, we obtain

$$\frac{pv}{T} \quad \frac{p_0 v_0}{T_0}.$$

The value of this quantity $\frac{pv}{T}$ is evi-

FIG. 221

dently proportional to the mass of gas in the cylinder, for it is clear that the volume would be twice as great if the mass of gas were to be doubled, keeping the temperature and pressure constant. Therefore if we let $\frac{pv}{T} = R$ when v is the volume of unit mass of the gas, we have

$$\frac{pv}{T} = mR$$

when v stands for a volume of the gas of mass m. The constant R depends on the *kind* of gas, since the volume of a unit mass of different gases is not the same. The most convenient form of this formula for use is

$$\frac{pv}{mT} = R \quad \text{or} \quad \frac{pv}{mT} = \frac{p'v'}{m'T'}$$

or *the product of the pressure by the volume, divided by the mass and by the absolute temperature, is constant for a given kind of gas.*
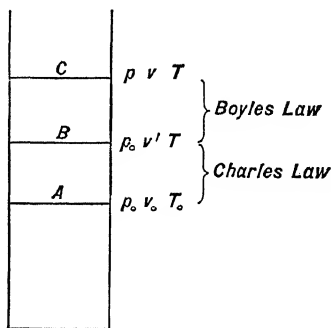
This formula is exact only to the degree that the gas obeys Boyle's law. It is, however, a very close approximation to the truth for the more perfect gases when far from their points of condensation.

A problem will illustrate the use of this formula. Suppose it is required to find the pressure that will be produced by 13 gms. of air in a vessel whose capacity is 1000 c.c. at 12° C. when it is known that 1 c.c. of air at 0° and 76 cms. pressure weighs 0.001293 gm. Substituting in the above formula we have,

$$\frac{p \times 1000}{13 \times (273 + 12)} = \frac{76 \times 1}{0.001293 \times (273 + 0)},$$

whence $p = 797$ cms. of mercury.

PROBLEMS

1. If a column of mercury at 100° C. and 90 cms. high balances the pressure of a column of mercury at 0° C. and 88.4 cms. high, find the density of mercury at 100°, that at 0° being 13.6.

2. Find the average coefficient of expansion of mercury between 0° and 100° from the data given in problem 1.

3. A barometer at 20° has a height of exactly 76 cms.; at what height would it stand if the mercury were at 0° C.?

4. A glass bulb has a capacity of 200 c.c. when placed in melting ice· How many grams of mercury will it contain at that temperature? What does the volume of the bulb become if placed in steam at 100°? Calculate the density of mercury at 100° and find how many grams will now be contained in the bulb. What weight of mercury flows out in the temperature change?

5. A glass bulb at 0° contains 544 gms. of mercury; what weight of mercury will flow out when it is heated to 90° C.?

6. How does the temperature of water affect the depth to which a hydrometer sinks in it?

7. A weighted glass bulb having a volume of 700 c.c. at 20° C. weighs 1.01 gms. when completely immersed in water at that temperature. What will it weigh in water at 4° C., taking the density of water as given in the table on page 158, under § 235, and taking 0.000025 as the volume coefficient of expansion of glass. *Ans. 0.05 gm.*

8. If 13 cu. ft. of air at pressure 76 and temperature 20° C. weighs 1 lb., find weight of 900 cu. ft. at 15° C. and pressure 55.

9. A mass of air at 100° C. and pressure 76 has a volume of 5 cubic meters; what will be its volume at 15° C. and pressure 90 cms.?

10. If a liter of air at 0° C. and pressure 76 weighs 1.293 gms., find how much a cubic meter will weigh at 25° C. and pressure 72.

CALORIMETRY

408. Basis of Heat Measurement. The quantitative measurement of heat rests on two assumptions. The first is that when bodies at different temperatures are put in contact the cooling of one and heating of the other are due to a transference of something which we call heat from one to the other, and that what one receives is precisely equivalent to what the other loses.

A second assumption is that heat is distributed uniformly throughout the mass of any homogeneous body all at the same temperature. That is, in a mass of water at one temperature every cubic centimeter contains as much heat as any other. These assumptions are justified by experiment, for the results of measurements based on them are so strikingly consistent that it seems almost as though we were dealing with a subtile substance. Indeed the chemists of 100 years ago were accustomed to think of heat as a substance and called it *caloric*. On account of this the process of measuring heat is called *calorimetry*.

Unit of Heat. Various units of heat are in use, but the one generally used in physical measurements is the heat required to raise the temperature of a gram of water one degree Centigrade. This unit is known as the *calorie*, but to be precise the exact temperatures must be specified. No general agreement has been reached on this point, but there are advantages in adopting as the unit *the heat required to raise the temperature of a gram of water from* 15° *to* 16° C., because 15° C. is somewhat near the average temperature at which experimental work is carried on, and this unit of heat is about equal to the one-hundredth part of the heat required to raise a gram of water from 0° C. to 100° C.

The following table shows the relative values of the calorie at different temperatures, taking that at 15° as the unit.

Temperature	Value of Calorie
5°	1.0049
10°	1.0021
15°	1.0000
20°	0.9982
25°	0.9973
30°	0.9971

In engineering practice the kilogram calorie or large calorie is used as a unit of heat on the continent of Europe, while in

English-speaking countries engineers usually employ the British thermal unit (written B. T. U.) which is the heat required to raise the temperature of a pound of water one degree Fahrenheit.

409. Specific Heat. It was discovered by the Scotch chemist Black (1728–1799) that the heat given out by a gram of lead in cooling one degree was by no means equal to that given out by a gram of iron when cooled the same amount, and that in general substances differed from each other in this respect.

The ratio of the heat given out by a mass of any substance in cooling one degree to the heat given out by an equal mass of water in cooling through the same range of temperature is known as the *specific heat* of the substance.

It may also be defined thus. *The specific heat of a substance is the number of calories required to raise the temperature of a gram of the substance one degree Centigrade.*

The following experiment illustrates how substances differ in their specific heats.

A number of balls of different metals, iron, zinc, copper, lead, and tin, of the same mass, are heated in a bath to a temperature of about 150° C. and then placed on a thin cake of paraffin supported above the table. The iron ball having the largest specific heat gives out the largest amount of heat in cooling and so melts the most paraffin. It therefore sinks deepest into the plate and perhaps drops clear through. The zinc and copper balls come next, while the lead having the smallest specific heat sinks in less than any of the others.

The specific heats of some substances are given in the table below.

TABLE OF SPECIFIC HEATS

Water at 4°..........	1.0049	Copper..............	0.0931
Water at 15°.........	1.0000	Zinc................	0.0935
Water at 30°.........	0.9971	Iron................	0.114
Ice at 0°............	0.502	Sulphur.............	0.176
Steam at 100°........	0.421	Aluminum............	0.217
Lead................	0.0310	Lithium.............	0.941
Mercury.............	0.0331	Crown glass.........	0.16
Tin.................	0.0562	Flint glass.........	0.117
Silver..............	0.0570	Normal ther. glass.....	0.199

It is remarkable that of all ordinary substances except hydrogen *water* has the greatest specific heat. The specific heat of a substance is in general greater at higher temperatures.

410. Calorimetry. The measurement of quantities of heat is called calorimetry. If the specific heat or gram calories of heat required to raise one gram of a substance one degree is represented by s, then the heat required to raise m grams of the substance one degree will be ms. And if the temperature is raised from t to t' the rise in temperature is $(t' - t)$ degrees, and the heat taken in by the substance is $ms(t' - t)$. This expression gives also the gram calories of heat given out when the substance cools through the same range of temperature.

Since the specific heat of a substance changes slightly with the temperature, s represents the *average* value of the specific heat between the temperatures t and t'.

The product ms is known as the *heat capacity* of the given mass of substance, *it is the number of calories of heat required to raise the whole mass one degree in temperature.*

Some methods of measuring quantities of heat are discussed in the following paragraphs, while certain other methods based on the melting of ice or the condensation of steam will be discussed later (§§ 437, 438, 453).

FIG. 222

411. Method of Mixtures. Suppose the specific heat of a mass of lead is to be measured by the method of mixtures. A weighed quantity of water is placed in a thin metallic cup or *calorimeter D* (Fig. 222) which is supported on little legs of cork or wood or some poor conductor of heat and is surrounded by an outside vessel to protect it from air currents and radiation from external objects. The mass of lead A having been heated to, say, 100° is suddenly plunged into the water and the water stirred till its temperature has risen as high as it will.
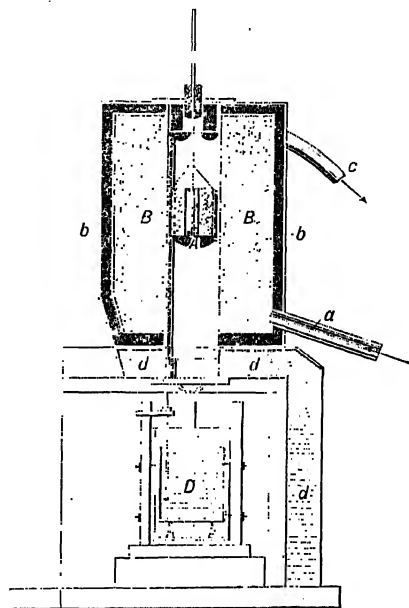
The heat given out by the lead in cooling from its initial temperature t' to the final temperature of the water t'', is $ms(t' - t'')$ where m is the mass of the lead and s is its specific heat. So also the heat taken in by the water as it warms from its original temperature t to t'' is expressed by $WS'(t'' - t)$, where W is the weight of water and S' is its specific heat, which in ordinary work is taken as 1.

But the heat given out by the lead must be equal to that received by the water, therefore

$$ms(t' - t'') = W(t'' - t).$$

In the above discussion the heat that went into the cup containing the water has been neglected. But clearly the cup must have experienced the same change in temperature as the water that it contains, and so must have received an amount of heat equal to $m's'(t'' - t)$, m' and s' representing its mass and specific heat. This heat also came from the lead and so must be added to the right-hand side of the above equation; the result is then

Heat given out by the lead in cooling		Heat received by water		Heat received by calorimeter vessel
$ms(t' - t'')$	$=$	$W(t'' - t)$	$+$	$m's'(t'' - t)$

therefore

$$s = \frac{(W + m's')\,(t'' - t)}{m(t' - t'')}$$

The quantity $m's'$ represents the quantity of heat measured in gram calories required to raise the temperature of the calorimeter cup one degree. It is called the *water equivalent* of the calorimeter because it represents the number of grams of water that would require as much heat to raise its temperature one degree as the calorimeter cup requires. It will be noticed that the water equivalent of the calorimeter is added directly to the mass of water which it contains. The water equivalent of the stirring rod and thermometer should be included also, as they, too, are raised in temperature by heat coming from the lead.

The form of heater shown in figure 222 was devised by Regnault. The substance to be heated is suspended in the central tube surrounded by the steam jacket. A thermometer with its bulb in a cavity in the middle of the substance serves to show

the steady temperature to which it finally comes. The calorimeter is slipped under the heater and the substance lowered into the calorimeter cup without exposure to cold air currents.

If the substance to be tested is in small fragments they may be held in a basket of light wire gauze whose heat capacity has been previously determined. If the solid is soluble in water some other liquid in which it does not dissolve must be used in the calorimeter. The specific heat of liquids as well as solids may be found in this way, provided there is no chemical action between the liquid and the water in the calorimeter cup which would cause either a development or an absorption of heat.

412. Compensation Calorimeters. In certain cases, especially where the calorimeter is of necessity large, it is important that the temperature of the instrument may be kept constant so that its heat capacity or water equivalent need not be known.

In the Junkers calorimeter, for instance, used to measure the heat developed in the combustion of gas or oil, a stream of cold water flows through a long copper pipe which is coiled around the combustion chamber. When all comes to a steady state the heat removed by the stream of water per second must be just equal to the heat arising from the combustion in the same time. The temperature of the water is measured at the inlet and also at the outlet by delicate thermometers, and the gain in temperature multiplied by the number of grams of water flowing through per minute gives the heat carried away by the stream of water in that time.

413. Electrical Calorimeters. The specific heats of liquids may be compared by heating first one and then the other in a calorimeter vessel by means of a current of electricity passing through a coil of wire immersed in the liquid. If the heat developed per second by the electric current is just the same in one case as in the other, and if the masses of liquid used in the two cases are such that the temperature rises at exactly the same rate in both cases, then the heat capacities of the two liquid masses must be the same; that is,

$$m_1 s_1 = m_2 s_2$$

where m_1 and m_2 are the masses of the two liquids and s_1 and s_2 are their respective specific heats.

414. Two Specific Heats of Gases. The specific heat of a gas may be measured while its pressure is kept constant, or it may be measured when the gas is enclosed in a bulb and kept at constant volume. *Experiment shows that the specific heat at constant pressure is greater than the specific heat at constant volume,* and when we come to discuss the relation of heat to work we shall find why this is so. (§ 424.)

Regnault measured the specific heats of various gases at constant pressure by causing a stream of gas to flow first through a long copper tube coiled in a vessel of hot oil, and then through a copper tube coiled in the calorimeter vessel and surrounded by water. The gas was heated by the oil bath and gave up its heat on passing through the calorimeter so that when the mass of gas which passed through the calorimeter was known its specific heat could be determined.

Great difficulty was found in measuring the specific heat of a gas at constant volume, because its heat capacity is very much less than that of the vessel in which it is enclosed; but the determination was successfully made by Joly using the steam calorimeter (§ 453).

SPECIFIC HEATS OF GASES

GAS	SPECIFIC HEATS		MOLECULAR WEIGHT	MOLECULAR HEAT	HEAT PER 1000 C.C. VOL. KEPT CONSTANT
	Constant Pressure	Constant Volume			
Air...........	0.237	0.169	...	....	0.222
Oxygen.......	0.217	0.155	32	4.96	0.221
Hydrogen.....	3.409	2.421	2	4.84	0.217
Nitrogen......	0.244	0.173	28	4.85	0.222
Carbon dioxide.	0.217	0.158	44	6.96	0.312
Chlorine.	0.121	0.096	71	6.81	0.427

The relative molecular weights of the gases are given in the fourth column. Evidently 32 grams of oxygen will contain the same number of molecules as 2 grams of hydrogen or 28 grams of nitrogen. The heat capacities of these weights of the various gases are shown in the next column, which indicates that the more perfect gases require nearly the same amount of heat per molecule to raise their temperatures one degree.

In the last column is shown the heat required to raise *equal*

volumes of the different gases one degree in temperature, a volume of 1000 c.c. at 0° C. and at atmospheric pressure being taken in each case. Here again is to be noted the equality of the values for the more perfect gases, as would be expected from Avogadro's law that *equal volumes of gases at the same temperature and pressure contain equal numbers of molecules*.

415. Change of Specific Heats with Temperature. The specific heats of the more perfect gases are nearly constant. The specific heats of solids and liquids are in general greater at high temperatures than at low. In case of most metals the change is small, but carbon, boron, and silicon show marked increase. These substances have been studied by H. F. Weber, who finds for the diamond at −50° specific heat 0.0635, while at 985° it is 0.4589, the value changing very rapidly at low temperatures and becoming almost constant about 800°. Also the various forms of carbon, graphite, and diamond differ greatly in their specific heats at low temperatures, but come to nearly the same value as the temperature is raised.

Nernst finds that the specific heats of substances at very low temperatures are approximately proportional to T^3, where T is the absolute temperature, but as the temperature rises they approach a maximum limit which nearly agrees with Dulong and Petit's law.

416. Dulong and Petit's Law. Dulong and Petit in 1819 showed that *the product of the specific heats of elements in the solid state by their atomic weights was approximately constant.* This constant is proportional to the heat required to raise one atom one degree and is therefore known as the *atomic heat.* Certain marked exceptions are boron, carbon, and silicon, but all of these substances have specific heats which vary greatly with the temperature, and are marked by particularly high melting points.

The table on page 284 shows the atomic heats in case of some substances.

PROBLEMS

1. When 10 lbs. of water at 12° C. is mixed with 17 lbs. of water at 20° C. find the temperature of the mixture.

2. If 10 lbs. of water at 12° C. is mixed with 18 lbs. of mercury at 20° C. what will be the temperature of the mixture?

3. If 3 kgms. of copper at 100° C. placed in 3 kgms. of water at 10° C. raise the temperature of the water to 17.7° C., find the specific heat of the copper.

Substance	Atomic Weight	Specific Heat	Atomic Heat
Aluminum..................	27.4	0.2170	5.94
Bismuth...................	210.0	0.0308	6.47
Cobalt....................	58.8	0.1067	6.27
Copper...................	63.4	0.0931	5.90
Iron......................	56.0	0.1138	6.37
Iodine....................	127.0	0.0541	6.87
Lithium...................	7.0	0.9408	6.59
Manganese................	55.0	0.1217	6.69
Lead.....................	207.0	0.0310	6.42
Platinum..................	197.4	0.0325	6.42
Silver....................	108.0	0.0570	6.16
Sulphur...................	32.0	0.1776	5.68
Tin......................	118.0	0.0548	6.46
Zinc.....................	65.2	0.0936	6.10
Boron (amorphous)........	10.9	0.254	2.77
Carbon (graphite)..........	12.0	0.174	2.09
Carbon (diamond).........	12.0	0.147	1.76
Silicon (crystalline).........	28.0	0.165	4.62

4. A mass of 300 gms. of platinum heated to the temperature of a furnace is dropped into 1000 gms. of water and raises its temperature from 15° C. to 25° C. Find the temperature of the furnace, taking the average specific heat of the platinum as 0.033.

5. A mass of 150 gms. of copper heated to 100° is dropped into 350 gms. of water at 12° contained in a thin copper vessel weighing 30 gms. Find the resulting temperature, taking the specific heat of copper as 0.094.

6. How much heat is required to warm the air in a room 3 × 6 × 5 meters in size, from 0° C. to 20° C., the pressure being constant?

7. A mass of 750 gms. of iron at 100° C. is dropped into a copper calorimeter containing 557.8 gms. of water at 15° C. and warms it up to 25° C. Find the specific heat of the iron if the copper vessel weighs 50 gms.

8. When 150 gms. of copper at 80° C. and 200 gms. of iron at 100° C. are dropped into 400 gms. of water at 12° C. contained in a copper calorimeter weighing 50 gms. find the resulting temperature.

9. How many calories in one British thermal unit (B. T. U.)?

Sources and Mechanical Equivalent of Heat

417. Sources of Heat. There are three principal sources of heat — chemical action, electric currents, and mechanical work. When two substances combine chemically the process is usually

accompanied by a giving out of heat. When the action goes on slowly, as when iron oxidizes or rusts, there is but slight rise in temperature, though the actual heat developed is the same as when the same amount of iron is burned in oxygen.

In ordinary combustion there is a rapid combination of the burning substance with the oxygen of the air. Heat must be supplied to start the process, but once started the heat of combination is sufficient to maintain it. Combustion may take place without the presence of air or oxygen. Copper and other metals will burn in chlorine gas. Gunpowder and other explosives contain within themselves all the elements which are to form the new combinations, so that when the spark or jar comes which precipitates the change it goes on with a rapidity which is explosive.

The development of heat by electric currents has become familiar to everyone in arc and incandescent lights and in the various industrial processes which use this source of heat. The laws which govern the heating effect of currents will come into our later study. But it should be noticed here that electric currents are always produced either by chemical action, by mechanical work, as in case of a dynamo driven by an engine, or by the direct action of some other source of heat, as in case of thermo-electric currents. So that in heating by electricity the ultimate source of the heat is the chemical action in the battery cells or the work done by the engine or water power driving the dynamo.

418. Heat of Combustion. The following table shows the heat developed in the combustion of a gram of various fuels.

HEATS OF COMBUSTION IN CALORIES PER GRAM

Hydrogen gas	34,500	Wood	4,000
Anthracite	7,800	Charcoal	8,000
Alcohol (absolute)	7,180	Gasoline	12,000

419. Heat Produced by Work. Heat may also be developed in a variety of ways from mechanical work, either against friction or in distorting viscous or plastic bodies or in compressing gases.

The savage obtains a fire by twirling, by means of a bow, a pointed stick pressed into a socket where it is surrounded by inflammable material. Everyone is now familiar with the great heat developed by the brakes on car wheels or in imperfectly

lubricated bearings, a " hot box " on a railway car often causing a blaze.

Those who held that heat was a *substance* — caloric — in order to explain the production of heat by friction, held that the frictional rubbing of substances caused some " latent heat " to become " sensible."

But Sir Humphrey Davy in 1799 caused two pieces of ice to be rubbed together by clock work in a vacuum thereby melting some of the ice, which showed that the current explanation was untenable since it was known that ice in melting *takes in* heat instead of giving it out.

In 1798 Count Rumford, who was in charge of the Bavarian cannon shops, being struck by the great development of heat in turning and boring cannon, caused a blunt boring tool to be turned when pressing into a socket in a metal block immersed in water. In this way water was made to boil and *it was shown that heat was developed so long as work was expended in driving the tool.* This experiment showed that heat could not be a substance forced out of the metal by the action of the boring tool, and Count Rumford remarks, " it appears to me extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner the heat was excited and communicated in these experiments, *except it be motion."*
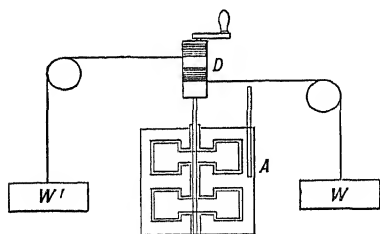
FIG. 223. Joule's mechanical equivalent apparatus

420. Mechanical Equivalent of Heat. The honor of establishing the equivalence of heat and work on a solid basis of experiment must be given to the English physicist, James Prescott Joule, who, in a series of careful experiments conducted between 1843 and 1850, measured by a variety of methods the amount of work required to heat a pound of water one degree Fahrenheit. In some of these experiments heat was developed by churning water, in others by churning mercury or by rubbing two plates of iron together, or by compressing air, or in rotating a bar of iron between the poles of a powerful

magnet, in which case the iron is heated by electric currents developed within it. *And all these diverse methods led to the same result, namely, that the energy required to heat one kilogram of water one degree Centigrade is equal to the work done in raising a weight of one kilogram to a height of 427 meters, or, in other units, 778 foot-*
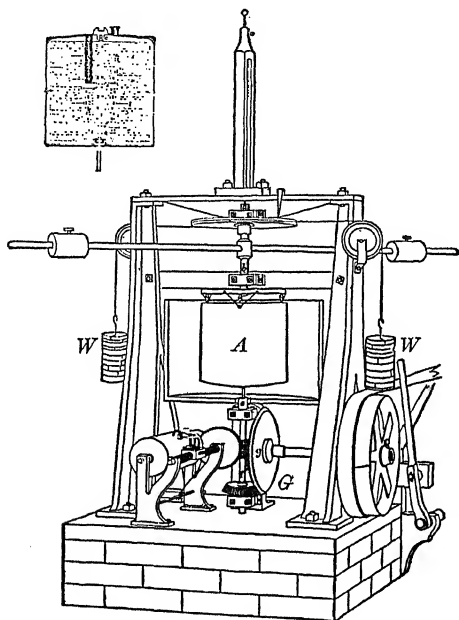
FIG. 224.   Rowland's apparatus for measuring mechanical equivalent of heat

pounds of work are required to raise the temperature of one pound of water one degree Fahrenheit.

For the most exact determination of the relation between heat and work Joule adopted the following method.

A closed calorimeter A, filled with water, was provided with a set of paddles attached to a central axle which could be rotated by means of the weights WW' which were suspended from cords wound around the axle D.   Fixed vanes projected inward from the sides of the calorimeter vessel so that between the fixed and rotating paddles the water was violently stirred.

The temperature of the water having been taken, the weights were wound up to their full height and then allowed to drop to the floor, turning the paddles as they descended. This was repeated twenty times and the temperature of the calorimeter again read.

The total work done was found by multiplying the amount of the weights by the distance through which they fell; but since the weights have some energy of motion when they reach the bottom this as well as the energy required to overcome the friction of the pulleys must be subtracted from the total work in order to obtain that spent in heat in the calorimeter.

A modification of this method used by Rowland made it possible to stir the water continuously, and at the same time measure the work. In this apparatus which is shown in figure 224, the calorimeter is suspended by a steel wire, while the shaft driving the paddles enters it from below. In making an experiment the paddles are driven at a uniform rate by an engine, and the tendency of the calorimeter to turn is exactly balanced by the weights WW, which are hung from cords attached to the rim of a wheel which is fastened to the calorimeter.

The work done is $2\pi nL$, where n is the number of revolutions of the paddles and L is the moment of the force exerted by the weights WW in balancing the calorimeter. The calorimeter was enclosed in an outer vessel to protect it from air currents and to enable the loss due to radiation to be accurately determined.

The mechanical equivalent of heat as determined by such experiments is found to be as follows:

1 *gram calorie* = 41,870,000 *ergs or* 4.187 $\times$ 10^7 *ergs.*
1 *kilogram   calorie* = 427.3   *kilogram-meters   of   work,   taking* $g = 980$.
1 *British thermal unit* = 778 *foot-pounds of work.*

421. Transformation of Heat into Work. The experiments of Joule showed that whenever mechanical work is apparently lost through friction, there is always a precisely equivalent amount of heat developed. It remained to show that whenever work is obtained from heat, as in any form of heat engine, *a quantity of heat disappears* which is equivalent to the work done by the engine. This was experimentally proved by the French

physicist and engineer, Hirn, who in an elaborate series of experiments showed that when an engine was doing work the total heat given out by it in the escaping steam, together with that lost by radiation and conduction, was less than that which it received from the boiler; and that *for every* 427 *kilogram-meters of work done by the engine, enough heat disappeared to raise the temperature of a kilogram of water one degree Centigrade.*

We therefore conclude that *heat is a form of energy* and that when work is done against friction there is a transference of energy, but no loss of it, as truly as in all other cases of work.

422. Heating of Gases by Compression. When a mass of gas contained in a cylinder is expanded by drawing out the piston the gas exerts a pressure against the piston as it moves outward and consequently does *work*. But in doing work it expends energy and consequently is *cooled*. On the other hand, if a mass of gas is compressed the work of compression is done *upon* the gas and its energy is correspondingly increased, and it is *heated*.

The heating effect of compression may be readily shown by the experiment of the *fire syringe*. This instrument consists of a strong glass tube or syringe, closed at one end and provided with a close-fitting piston. If the tube is full of air and the piston is suddenly

FIG. 225. Fire syringe

thrust home, the heat developed will be sufficient to ignite a bit of tinder. Instead of tinder a pellet of cotton soaked in ether may be used, in which case the flash is readily seen as the piston is forced down.
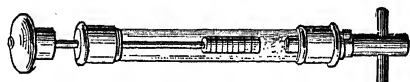
This experiment shows that the dynamical heating of a gas when compressed is very considerable. When air at 0° C. is compressed in a non-conducting cylinder, its rise in temperature is 90° when compressed to half its original volume, 429° when compressed to one-tenth, and 1084° when compressed to one-fiftieth of its volume.

In air compressors the heat developed in this way has to be removed by a stream of cold water.

423. Cooling Due to Work of Expansion. A gas when it expands is cooled because it does work; but *is all of the cooling due to the external work done?*

That is, *if it were possible to pull out the piston of a cylinder containing gas so suddenly that the gas could not follow it and exert pressure against it as it moved back, would the gas be cooled or not?*

This question was asked by Joule, and answered by an ingenious experiment in which he connected a copper receiver containing air at a pressure of 22 atmospheres, with another from which the air had been exhausted. On opening the stopcock between the two vessels the air expanded and filled both, but of course *it did no external work in expansion* since there was no piston to push back.

It was found after the expansion that the mass of gas as a whole had not changed in temperature, the gas rushing into the vacuum being heated just as much as the expanding gas in the other vessel was cooled.

More exact determinations show, however, that most gases when expanded are *slightly* cooled even when no external work is done and it is this cooling of which advantage is taken in the process of making liquid air (§ 463). This is because of a small amount of energy required to overcome a slight attraction which exists between the molecules at their distances of separation in the gaseous state.

424. Specific Heats of a Gas. We can now see why the specific heat of a gas at constant pressure must be greater than that at constant volume. For when a mass of gas is warmed while the pressure is kept constant, it *expands*, doing external work. *The heat supplied must, therefore, furnish the energy for this work* as well as that which simply increases the energy of motion of the gas molecules. But when a gas is kept at constant volume there is no external work done and the heat supplied all goes to increase the molecular energy of the gas. According to Joule's experiment (§ 423), the increase in molecular energy is just the same in one case as in the other, so that *the difference between the heats required in the two cases is entirely due to the external work done, and is mechanically equivalent to that work.*

425. Convective Temperature of the Atmosphere. The change of temperature caused by the compression or expansion of air plays a most important part in the atmosphere. Masses of air moving upward expand and cool, while descending air masses are heated by compression. This in part serves to determine the

distribution of temperature in the atmosphere, the temperature at any height tending to be equal to that which a mass of air rising to that point from the surface of the earth acquires in consequence of its expansion.

The presence of water vapor modifies what may be called the convective temperature at a given height, for the latent heat given out as the moisture in a rising mass of air condenses retards the cooling. A mass of dry air at 20° C. at the earth's surface, will be cooled to $-53°$ C. in rising $3\frac{1}{2}$ miles.

When there is a downward current of air, as in case of the wind blowing over mountains and sweeping down into the valleys beyond, the compression of the air as it descends raises its temperature so that it becomes a warm wind, as in the so-called "foehn" wind of the Alps or the "dry chinook" of Montana.

426. The Nature of Heat Energy. When a body is heated it radiates heat to surrounding bodies. The rate at which a given body gives off radiation depends on its temperature. As it grows hotter the radiation may become so intense that the body glows or is incandescent. Later it will be shown that this radiation is made up of waves in which the vibrations are almost inconceivably rapid. These waves originate in the hot body and take energy from it so that it cools as it radiates. It is believed that these waves are a consequence of rapid vibratory motions in the molecules of the body, and that the heat energy of a body exists, in part at least, in the form of energy of motion of the molecules.

Radiation comes from all bodies even those that we ordinarily speak of as cold. The molecules of all bodies are therefore considered to be in rapid vibration, though we are ignorant of the exact nature of this vibration.

But heat energy exists in bodies in another form than energy of vibratory motion, for bodies usually expand when heated, and consequently the particles or molecules are slightly moved apart. And since in case of solids and liquids there is a strong attraction between the particles, work must be done in separating them, and the energy which does this work comes from the vibratory energy of the molecules, which is thus transformed and stored up in the body as potential energy.

Another instance of such a transformation is in case of *change*

of state, as when ice is melted. Here also particles which are held firmly in a comparatively fixed position in the solid state are dragged away from each other and set free to slip past each other in the liquid state. To effect this change work must be done, and consequently in this case also a certain amount of vibratory heat energy must be changed into energy of separation or potential energy.

Experiment is in complete agreement with this conclusion and shows that a considerable amount of heat energy must be given to a body to change it from the solid to the liquid form.

Therefore, heat energy is thought of as existing in the body both in the form of kinetic energy or energy of motion of molecules, and as potential energy due to the separation of molecules in opposition to their mutual attractions.

427. Temperature Depends On the Kinetic Energy of the Molecules. When two bodies at different temperatures are put in contact there is a transfer of molecular energy from one to the other until equilibrium is established. When there is no longer any change taking place, it is said that both are at the same temperature. This transfer of energy is doubtless due chiefly to the energy of motion of the molecules, as it is difficult to see how the potential energy of the molecules of one body could affect appreciably the condition of a neighboring body. When ice and water are mixed together until both come to the same temperature, all flow of heat from one to the other entirely ceases and yet a gram of ice has very much less potential energy than a gram of water at the same temperature. It appears, therefore, that *temperature* is chiefly, if not entirely, determined by the energy of motion, rather than the potential energy, of the molecules.

PROBLEMS

1. The cylinder of an air compressor is cooled by a stream of water in which the flow is 1 gallon per minute. If 10 H.P. is expended in compression, find how many degrees the water is raised in temperature. 1 gallon = 3785 c.c. 1 H.P. = 746 × 10^7 ergs per sec.

2. How much is the water of Niagara raised in temperature by the fall of 160 ft.?

3. What would have to be the velocity of a lead bullet that it may be melted on striking the target, supposing all its energy to be transformed

into heat within the bullet? Assume the temperature of the bullet to be the melting point of lead. It takes 5.86 calories to melt 1 gm. of lead.

4. What is the heat of combustion of anthracite coal in British thermal units per pound?

5. How much more heat is required to raise the temperature of a kilogram of air from 0° C. to 30° C. constant pressure, than if the volume were kept constant, and why is more required?

6. How many British thermal units of heat are developed by the brakes when a 100-ton train having a velocity of 20 miles per hour is brought to rest?

7. One liter of air at 0° C. is warmed to 10° C. at constant pressure. Compute the amount of heat required, and the external work done by its expansion, taking the pressure as 76 cms.

Also compute the heat that would have been required if its volume had been kept constant.

From these two results deduce the mechanical equivalent of heat (see § 424).

TRANSMISSION OF HEAT

428. Different Modes. Three modes of transferring heat energy from one place to another are recognized, *conduction, convection,* and *radiation.*

When heat energy gradually diffuses through a mass of matter, passing from particle to particle from the warmer toward the colder parts of a body, the process is called conduction. In this case energy of motion is conceived as communicated from molecule to molecule progressively throughout the mass.

When heat is carried along by the motion of a stream of gas or liquid, the process is called convection.

In the above two cases the transference takes place in and through *matter,* but a hot body surrounded by a perfect vacuum may give out energy and warm neighboring objects. *In this case the energy is transmitted by waves in the ether and the process is called radiation.* The term radiation is also applied to the ether waves themselves coming from the hot body.

When one end of a bar of iron is heated the other end becomes hot by conduction; the circulation in a vessel of water which is being heated carries heat from one part to another by convection; while the warmth received from hot coals in an open fireplace comes to us largely as radiation.

Conduction and convection are relatively slow processes while radiation is transmitted with the speed of light.

In transparent bodies, such as glass or water, heat is communicated from one part of the substance to another by conduction and radiation combined, for energy is radiated through the body directly from one part to another at the same time that it is being communicated from molecule to molecule by conduction.

While radiation originates in hot bodies and heats any body which absorbs it, *radiation itself cannot be regarded as heat; for unless it is absorbed it does not affect the temperature of the bodies through which it passes.* We shall study the nature of radiation in connection with *light;* while radiation considered as an *effect* of heat will be discussed in §§ 473–485.

429. Conduction. In general solids conduct heat better than liquids, and liquids than gases. Silver and copper are the best
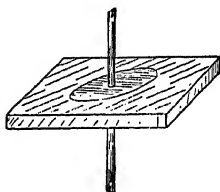
FIG. 226
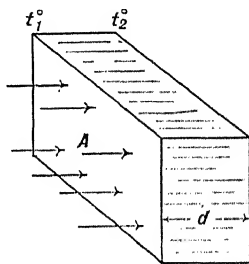
FIG. 227

conductors of heat, having about **7** times the conducting power of iron, while iron conducts **100** times as well as water, and water has **25** times the conductivity of air.

Among solids the metals are the best conductors, and it is remarkable that, *generally speaking, the best conductors of heat are also the best conductors of electricity.*

In crystals heat may be conducted more rapidly in one direction than another. If a thin plate of quartz is coated with wax or paraffin and if a wire kept hot by an electric current is passed through a hole in the center of the plate, the wax will melt outward in elliptical form if the plate is cut parallel to the axis of

the crystal, showing that heat is conducted more rapidly in the direction of the axis than at right angles to it.

430. Conductivity. If a slab of substance of uniform thickness d and faces of area A has one surface at temperature t_1, and the other at t_2, the heat H which is transmitted per second will be proportional to the area of the faces A and to the difference in temperature of the surfaces $t_1 - t_2$, while it will be inversely proportional to d, the thickness of the plate. Thus,

$$H = \frac{kA(t_1 - t_2)}{d}$$

where k is a constant which depends on the substance of which the slab is made, and is known as its coefficient of conductivity or simply its *conductivity*. When A is one square centimeter and d is one centimeter, and $t_1 - t_2$ is one degree, then $H = k$; that is, *the conductivity or conducting power of a substance is measured by the number of gram calories of heat which are transmitted in one second through a plate one centimeter thick and having surfaces one square centimeter in area when the opposite faces differ in temperature by one degree Centigrade.*

The drop in temperature per centimeter between one side of the plate and the other is called the *temperature gradient* and is expressed in the above formula by

$$\frac{t_1 - t_2}{d}.$$

431. Measurement of Conductivity. The conductivity of a metal which is a good conductor may be measured by the following method due to Searle. A short bar, say 1 in. in diameter and 6 in. long, is heated at one end by steam while the other end is kept cool by a stream of water which flows through a pipe closely wound around the bar in a helix as shown in the figure. The whole is thoroughly packed in hair felt, which is a very poor conductor, so that heat cannot escape from the sides of the bar but must flow from the hot toward the cold end, where it is removed by the stream of water. When a steady condition of flow is reached, the heat passing along the bar in any time will be equal to that removed in the same time by the stream of water. Thermometers are mounted at the ends of the helical tube, by

which the temperature of the water may be observed as it flows in and flows out, and the gain in temperature of the water as it flows through the helix multiplied by the total weight of water that passes through in, say, 10 minutes, gives the gram calories of heat transmitted along the bar in that time. The temperatures at two points on the bar, such as t_1 and t_2, are measured by
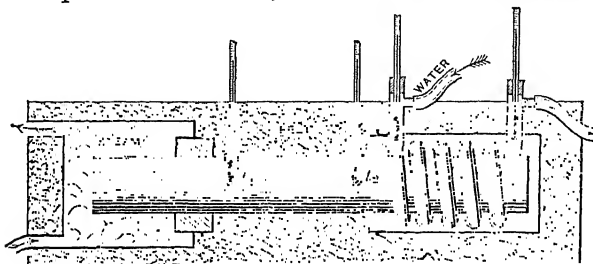
FIG. 228

thermal junctions or thermometers, the distance d between those points is measured and also the area of cross section A of the bar. All the elements are then known for computing the conductivity from the formula of the preceding article.

In case of poor conductors the flow of heat through a thin slab may be measured by such a method as that of Lees. A very broad and thin disc-shaped box of copper, heated by means of a current of electricity which passes through a resistance coil contained in the box, is placed between two thin plates of the substance of which the conductivity is to be determined, say these are plates of glass. Outside of the glass plates are placed solid copper discs, so that each plate of glass comes between the central heating box and an outer disc, as shown in section in figure 229. The temperatures of the outer discs and also of the inner box are determined by thermometers or thermal junctions (§ 682), and since copper is so good a conductor compared with glass the surfaces of the glass plates may be supposed to be at the same temperatures as the copper plates with which they are in contact. A
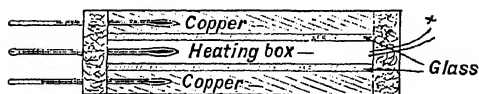
FIG. 229

band of thick felt encircles the edge of the discs to prevent loss in that way and the heating is maintained constant until a steady state is reached where there is no longer any change in the temperatures and the heat flows out through the glass plates as fast as it is generated.

Then the heat developed per second is easily determined by electrical measurements and knowing the temperatures of the surfaces of the glass plates and their areas and thickness the conductivity of glass may be calculated.

432. Conductivity of Liquids. The conductivity of liquids is small, that of mercury being only one-tenth that of iron, and the determination of their conductivities is complicated by convection currents. They may be determined by the method described in the last article. A thin layer of liquid rests on a copper disc and has the copper heating box in contact with its upper surface. By this arrangement convection currents are not established, and if suitable precautions are taken to prevent loss of heat upward from the heating box, the conductivity of the liquid may be measured.

433. Conductivity of Gases. The conductivity of a gas is measured by the rate of cooling of a heated bulb enclosed in a spherical vessel of the gas to be studied, the outer surface of the large vessel being kept cool in a water-bath. The determination is complicated both by convection and radiation.

But the kinetic theory shows that the conductivity of a gas should be the same at low pressures as at high, provided the rarefaction is not so great as to make the mean free path of the molecules appreciably large compared with the size of the vessel.

By diminishing the pressure of the enclosed gas the effect of convection is diminished, while the true conductivity is not affected. In this way the two may be separated. To determine the effect of radiation the gas is exhausted as far as possible, so as to do away with both convection and conduction.

TABLE OF CONDUCTIVITIES

Heat Conductivities in C. G. S. Centigrade Units

Silver.......	1.096	Paraffin.....	0.0002	Air...........	0.000056
Copper......	1.000	Hair felt....	0.00009	Hydrogen.....	0.000327
Iron........	0.167	Cork.......	0.0007	Carbon dioxide	0.000030
Zinc........	0.265	Water......	0.0014		
Glass.......	0.0020	Mercury....	0.0152		

PROBLEMS

1. A glass window pane 170 cms. long and 90 cms. wide is 3 mm. thick. How much coal must be burned per hour to compensate for the loss of heat by conduction when the outer surface is at $-5°$ C. and the inner surface at 20° C.?

2. In the preceding problem if the temperatures given are those of the outer air and within the room, respectively, will the flow of heat be as great as found above? Explain answer.

3. A partition of iron 2 cms. thick and 10 cms. high and 15 cms. wide divides a vessel into two compartments, one of which contains ice, while steam at 100° is passed into the other. Find how much ice is melted in 5 minutes, when 80 calories are required to melt 1 gram of ice.

4. An iron boiler has 1 square meter of heating surface. How much water will be evaporated in 1 hour when the outer surface is kept at 150° C., while water is boiling at 100° C., if the iron is 0.7 cm. in thickness, and if 536 calories are required to evaporate 1 gram of water?

CHANGE OF STATE

Fusion

434. Changes of State. Among the most interesting and important effects of heat are the *changes of state* which it produces in matter. Solids if sufficiently heated are changed to liquids, or may pass directly into the gaseous state, and liquids are transformed into vapors or gases. Even with the temperatures that can be artificially produced almost all known substances can be made to assume any one of these three conditions. Thus nitrogen may be liquefied and solidified, while, on the other hand, platinum may be liquefied and volatilized.

There are three principal changes to be considered: that from the solid to the liquid state, that from liquid to gas or vapor, and that directly from solid to vapor or gas.

Melting. When ice below zero is slowly heated it first warms to 0° and then melts, the temperature at the surface of the ice remaining constant at 0° until it is all melted. If heat is now slowly taken from the mass, solidification will take place at the same temperature and it will remain at 0° till all is frozen. That the temperature should remain constant while the substance is melting, although it is steadily receiving heat, is characteristic of

all cases where there is a well-marked melting point. Some substances, such as selenium, pass from the solid to the liquid state through a soft pasty condition and without there being any point at which the temperature is stationary. Iron passes through such a pasty stage, on which account it is easily welded. So also glass, beeswax and paraffin become very soft as the melting point is approached. There is, however, usually some temperature at which the pasty mass becomes fluid and a considerable absorption of heat takes place, and this is the melting point.

MELTING POINTS

Tungsten....	3400° C.	Silver.........	955° C.	Tin.........	232° C.
Platinum.....	1710°	Aluminum....	657°	Sulphur.....	114°
Iron.........	1503°	Zinc..........	419°	Ice.........	0°
Copper......	1084°	Lead.........	327°	Mercury....	−39°

Change of Volume. Most substances occupy a larger volume in the liquid state than in the solid, and therefore contract when they solidify. Some, however, among which are water, bismuth, and cast iron, expand when they solidify. This property is of importance in making castings, as all parts of the mould are filled and its details are sharply reproduced in the casting. The volume of 1 gram of ice at 0° is 1.09082 c.c., while that of the same amount of water at 0° is 1.00012 c.c., so that the increase in volume of a cubic centimeter of water on freezing is 0.0907 c.c.

In consequence of the expansion of water in freezing, ice floats in fresh water with about one-twelfth of its volume exposed. The practical importance of this property of water can hardly be overestimated. If it contracted in freezing, ice would sink to the bottom of a lake, and ultimately the whole mass of water would be frozen solid, to the destruction of all forms of life that it contained.

435. Effect of Pressure on Melting. *If a substance contracts when it melts, increase in pressure will aid melting; that is, the melting will take place at a slightly lower temperature under increased pressure.* If, on the other hand, a substance expands when it melts, the effect of increased pressure will be to raise the melting point, as the pressure in a sense resists the melting.

In case of water an increase of pressure amount to 1 atmosphere

will lower the freezing point 0.0075° C. hence an increase of pressure of 133 atmospheres, or about 1900 lbs. to the square inch, will be required to lower the freezing point one degree.

In some experiments on the expansive force of water in freezing made by Major Williams at Quebec, iron bomb shells were filled with water and exposed to cold. In one case the plug was driven violently out and a short column of ice protruded from the opening. In another experiment the shell was burst and a sheet of ice was forced through the crack. In these cases probably the water was still unfrozen under the enormous pressure just before the shell burst, and froze instantly as the pressure was relieved by the bursting.
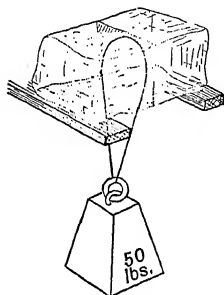
FIG. 230. Wire cutting through a block of ice

If a weight of 40 or 50 lbs. is suspended by a wire loop hung over a block of ice at 0° C., the wire will cut slowly through the ice, the pressure causing ice to melt under the wire; but the water flowing around the wire freezes again above it, leaving the block as solid as before. As this action goes on heat is taken up by the water in melting at the lower side of the wire and given out again in freezing on the upper side, so that there is a steady flow of heat across the wire as it cuts its way through the ice. And as the action cannot take place without this transfer of heat, a copper wire, being a good conductor of heat, will move faster than an iron wire, other things being the same.

Regelation, or the clinging together of pieces of ice when pressed together, is doubtless due to melting at the points of contact where there is pressure, followed by instant freezing when the pressure is relieved. In this way may be explained the packing of a snow ball, so also Tyndall explains the motions of glaciers; for where the ice is under special pressure, as where it meets a projecting point of rock, it melts and the water flowing around the obstacle freezes again.

436. Latent Heat of Fusion. If a vessel containing ice and water at 0° C. is kept in a region where everything is at 0°, there will be no change, the ice will not melt nor will the water freeze. But if the temperature of the surrounding bodies is below 0° there

will be a flow of heat out of the water accompanied by freezing; if, on the other hand, heat is given to the mass ice will melt, but the temperature *at the surface of contact of ice and water* will remain steady at 0° until all is either melted or frozen. It appears, then, that *a substance may be at the melting point, but it will not melt unless a definite amount of heat is received for each gram of substance melted.*

The latent heat of fusion is the quantity of heat required to change one gram of a substance from the solid to the liquid state without change of temperature.

If equal weights of water at 0° and at 100° are mixed together the temperature of the mixture will be 50°. But if equal weights of ice at 0° and water at 100° are mixed the temperature of the resulting mass of water is 10°.

Every gram of water cooling from 100° to 10° gives out 90 calories of heat, but for every such gram cooled from 100° to 10° there is a gram of ice melted into water at 0° and then raised from 0° to 10°. For the latter change just 10 calories is required, therefore 80 calories must have been used in transforming 1 gram of ice at 0° into water at 0°. The latent heat of fusion of ice is therefore 80.

Suppose m grams of ice at 0° C. are put into a calorimeter containing W grams of water at a temperature t and the temperature of the mixture after the ice is melted is t'; then, calling the latent heat of fusion of the ice L, the heat required to melt it to water at 0° will be Lm, and mt' calories more will be required to raise it to temperature t'. The water in the calorimeter will give out in cooling $W(t - t')$ calories, where W includes the water equivalent of the calorimeter. Then

$$Lm + mt' = W(t - t').$$

Of course if Wt is not as great as Lm, only a portion of the ice will be melted and the final temperature will be 0°.

The latent heat of fusion doubtless represents the energy required to separate the molecules from the close association which prevails in the solid state; it probably exists in the form of potential energy. So when a body is heated and *expands* it is no doubt true that only a part of the heat given to it contributes to its rise in temperature, the rest of the heat energy doing the work of expansion in opposition to the internal forces of cohesion and also to the external pressure. This portion of what is called the specific heat exists as potential energy and might appropriately be called the *latent heat of expansion.*

LATENT HEATS OF FUSION

Ice	80.00	Lead	5.36
Sulphur	9.37	Zinc	28.13
Tin	14.25	Silver	21.07
Bismuth	12.64	Mercury	2.82

437. Ice Calorimeter. Some important calorimetric processes make use of the latent heat of fusion.

The ice calorimeter of Lavoisier and Laplace is shown in the figure. An inner vessel to receive the heated substance is surrounded by ice in a vessel which is again surrounded on all sides by ice contained in an outer vessel. Heat from outside will melt ice in the outer vessel, but the ice in the inner vessel will melt only very slowly and its rate may be determined by the drip from the lower stopcock. When a mass of heated metal is introduced into the inner vessel it cools down to 0° and gives out heat which causes an increased flow from the lower stopcock. The water escaping from this, in excess of the steady drip observed at first, gives the weight of ice melted by the heat from the substance.
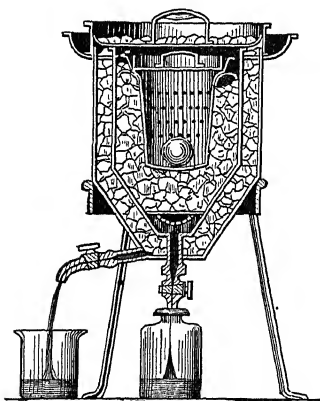
FIG. 231. Ice calorimeter

Let m be the mass of the substance which was heated to a temperature t, and introduced into the calorimeter; and let L be the latent heat of fusion of ice, and W the weight of ice melted; then if s is the specific heat of the substance

$$ms(t - 0°) = LW.$$

This method is only adapted to determine the specific heat of rather large masses, as the water from the melting ice clings to the fragments of ice and escapes only gradually.

438. Bunsen Ice Calorimeter. A form of ice calorimeter which can be used to measure small quantities of heat with much precision was devised by Bunsen and is represented in figure 232. It depends on the change in *volume* which takes place when ice

melts. The vessel PWQ is made of one piece of glass in the form of a bulb W enclosing a test-tube P, and provided with a curved neck Q. The bulb W is filled nearly to the bottom with distilled water, the remainder of the bulb and the neck Q being filled with mercury. The whole instrument is now carefully packed in snow and cooled to 0° C.; after which a freezing mixture is introduced into the test-tube P and a sheath of clear ice frozen around it.
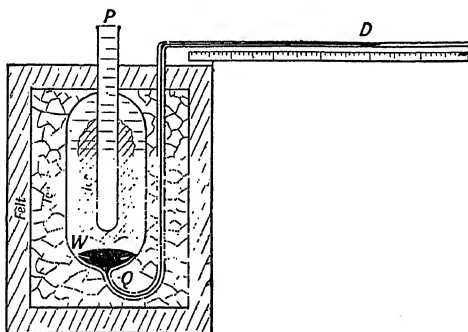
FIG. 232.   Bunsen ice calorimeter

When the ice is sufficiently thick the freezing mixture is removed and the test-tube filled with water at 0° C.

Attached to Q is a long capillary tube of uniform cross section extending horizontally. As the water in freezing expands it forces the mercury out to the end of this tube. Now let a fragment of substance heated to 100° be dropped into the test-tube: it gives up its heat to the water, and then to the ice surrounding the tube and causes some ice to melt. But in melting a contraction in volume takes place of 0.0907 c.c. per gram of ice melted; and the contraction may be determined by observing how far the mercury column has moved back along the tube D. From this contraction the weight of ice melted may be determined and so the heat given out by the substance in cooling from 100° to 0° C. becomes known. The volume of the contraction when the mercury moves back a certain distance along the tube D may be found by weighing the mercury required to fill a measured length of the tube.

439. Retardation of Freezing Point. Substances may often be cooled below the temperature at which they normally solidify, and still remain in

the liquid form. As soon as solidification begins, however, the temperature of the mass rises owing to the latent heat given out by the part that is becoming solid, and when the temperature has been raised in this way to the melting point no further solidification takes place.

440. Supersaturated Solutions. The formation of crystals from a concentrated solution is somewhat analogous to solidification. In many cases when a crystal is dissolved heat becomes latent, and when it forms from a solution latent heat is given out. Sodium sulphate crystals may be melted at 48° C. in their own water of crystallization. If the solution is now allowed to cool slowly in a clean flask closed by a cork, it may be brought down to 15° or 20° without crystallizing. Dropping in a minute crystal of the salt will at once precipitate the crystallization, which will go on so rapidly that the temperature may rise 10° or 15°, but not higher than 48° C., the rise in temperature being produced by the giving out of what may be called the latent heat of crystallization.

441. Freezing Mixtures. When a very dilute solution of common salt in water is cooled below 0° C., ice crystals are formed, leaving the remaining solution stronger; while if a saturated solution of salt is cooled in the same way, the salt crystallizes out, leaving the solution weaker. This goes on progressively as the temperature is still further lowered, one solution becoming stronger and the other weaker, until at −22° C. the two solutions reach the same strength, and when cooled further each solidifies into a mass which the microscope shows to be an agglomeration of minute crystals both of ice and of salt.

The final solution, which is of such strength that on cooling neither component crystallizes out without the other, is said to be *eutectic*.

Similarly, there may be a eutectic alloy of two metals, the melting point of which is a minimum for the given metals.

Now, suppose that a freezing mixture of common salt and ice at 0° C. is enclosed in a non-conducting vessel. The affinity of the two causes both salt to dissolve and ice to melt, and in each of these changes heat becomes *latent;* and this heat energy must come out of the mixture, which accordingly is cooled. But it appears from the first part of this article, that *ice and salt cannot both be in equilibrium with the same brine solution, unless it is a saturated solution at* −22° C. Consequently ice and salt continue to dissolve until this final state is reached. No lower temperature than this can be produced by a mixture of these substances. By using calcium chloride and ice, a temperature of −54° C. may be attained.

PROBLEMS

1. How much ice is melted when a mass of 500 gms. of copper at 100° C. is dropped into a hole in a block of ice at 0° C.?

2. How much more energy has 1 kgm. of water at 70° C. than the same mass of ice at −10° C.?

NOTE: The specific heat of ice is not the same as that of water.

3. A mass of ice weighing 30 gms. is in a tube with water enough to make the whole volume 50 c.c. at 0° C. What change in volume takes place when 100 gram-calories of heat are given to the mixture?

4. When 400 gms. of ice at 0° C. are put into 500 gms. of water at 60° C., what is the final temperature of the mixture?

5. If 5 lbs. of snow are mixed with 2 lbs. of water at 60° C., how much snow will be melted?

6. If 3 gms. of iron at 100° are dropped into a Bunsen ice calorimeter, find the resulting change in volume.

7. When 2 lbs. of snow at 0° are mixed with 3 lbs. of water at 80°, find the resulting temperature.

8. A mass of 100 gms. of ice at −16° C. is put into water at 0° C. and 10 gms. of the water are frozen, all coming to 0° C. Find the specific heat of ice.

9. If 100 gms. of lead cools from 340° to 327° in 2 minutes and then the temperature remains steady for 25.8 minutes while the mass is solidifying, find the latent heat of fusion of lead, assuming that heat is lost at a uniform rate and that the specific heat of lead at 330° is 0.032.

Vaporization

442. Evaporation. The change from the liquid to the gaseous state is known as *evaporation*, or if accompanied by the formation of bubbles of vapor throughout the mass it is called *boiling* or *ebullition*.

Ordinary open-air evaporation is complicated by the presence of the gaseous atmosphere above the surface of the liquid. The simple case where a vessel contains nothing but a liquid and its own vapor will first be considered.

443. Saturated Vapor. Take a barometer tube about a meter long, fill it carefully with mercury so as to exclude all air and invert it in a deep cistern of mercury (Fig. 233). The mercury in the tube will stand at the barometric height if the tube is raised high enough. Now introduce into the tube a few drops of ether. On reaching the vacuum the ether will at once evaporate and the mercury column will be forced down perhaps 40 cms. by the pressure of the vapor. If there is enough ether all will not evaporate, but a little will remain as a liquid on top of the mercury column; *in this case the vapor is said to be saturated, and its pressure is the greatest possible for ether vapor at the given temperature.* For if the volume occupied by the vapor is dimin-

ished by pushing the tube downward some of the ether vapor will condense, but the height of the mercury column will remain unchanged, showing that there has been no change in pressure. So also if the tube is raised, thus increasing the volume of the vapor, ether evaporates, but the pressure does not change until all the ether is evaporated.

A saturated vapor is one which is so dense that it cannot be further compressed without condensation. If the volume of a saturated vapor is diminished, an exactly corresponding amount of vapor condenses, so that the pressure and density of the remaining vapor continue unchanged so long as the temperature is constant.

FIG. 233. Pressure of ether vapor

A saturated vapor is in equilibrium with its liquid, the tendency of the liquid to evaporate being exactly balanced by the tendency of the vapor to condense. Indeed it is probable that molecules are constantly escaping from the liquid and passing into the vapor, while other molecules of vapor striking down into the liquid are caught and held by its attraction, and *when the vapor is saturated these two processes exactly balance.* (See § 447.)

FIG. 234. Measuring vapor pressure

444. Non-saturated Vapor. When a vessel containing liquid and vapor is enlarged so much that all the liquid is evaporated, a further enlargement of the vessel causes the pressure of the vapor to diminish very nearly according to Boyle's law for gases, and the more it is expanded and so removed from its point of condensation, the more exactly does it conform to Boyle's law.

445. Influence of Temperature on the Pressure of Saturated Vapors. The series of changes considered in § 443 is supposed to have taken place at a constant temperature. *The pressure*

of a saturated vapor increases as the temperature rises. The effect of temperature on vapor pressure may be determined by the apparatus of figure 234. Two barometer tubes are surrounded by a water bath whose temperature may be varied. A few drops of the liquid to be studied are introduced into one of the tubes and float on the mercury column, filling the upper part with vapor, the pressure of which causes the mercury to stand lower than in the other barometer. The difference in height of the two mercury columns thus measures the pressure of the vapor. Evidently this method can be applied only when the pressure of the vapor is less than one atmosphere.

The following tables give the vapor pressure of water and of a few other liquids.

VAPOR PRESSURE OF WATER

TEMPERATURE	PRESS. MM. OF MERCURY	TEMPERATURE	PRESS. MM. OF MERCURY
−10° C.	2.16	99.9° C.	757.30
0	4.58	100.0	760.00
+10	9.18	100.1	762.71
20	17.41		
30	31.56	TEMPERATURE	PRESSURE IN LBS. PER SQ. INCH
40	55.0		
50	92.2	100° C.	14.7
60	149.2	110	20.8
70	233.8	150	69.1
80	355.5	200	225
90	526.0	250	576
100	760.0	Subtract 14.7 from above to get steam-gauge pressure.	

VAPOR PRESSURES OF SOME OTHER LIQUIDS

TEMPERATURE	ALCOHOL	ETHER	MERCURY
−20° C.	0.334 cm.	6.92 cm.	
0	1.27	18.23	0.00002 cm.
20	4.40	43.48	0.0001
50	22.03	126.8	0.0015
100	168.5	492.0	0.027
Boiling points...	78°	34.6°	357°

Notice how small the pressure of mercury vapor in a barometer must be.

446. Density of Saturated Vapor. The higher the temperature of a given mass of vapor, the smaller the volume into which it must be compressed before condensation begins.

The density of a saturated vapor therefore increases with rise in temperature.

447. Evaporation in Air. If water is introduced into a large vessel full of air, very nearly the same amount will evaporate as if the air had not been there. The water vapor exerts its own pressure independent of that of the air, making the total pressure the sum of the two.

This is a particular case of the general law stated by Dalton as follows: *when a liquid is contained in a vessel with air or any other gas or vapor that has no chemical action upon it, the amount that will evaporate and the pressure of its saturated vapor will be the same as though the other gas were not there, and the total pressure will be the sum of the pressures of the gas and of the saturated vapor.*

The law thus stated is a close approximation to the truth, but it does not hold *exactly*, especially when the gas or vapor is very dense.

The presence of air or gas has a very marked effect, however, in *retarding* evaporation. The layer of air next the liquid becomes filled with vapor which is gradually carried away by currents and by diffusion and as it is removed further evaporation takes place till saturated vapor fills the vessel. The fact that the *pressure of another gas or vapor does not stop the evaporation of a liquid,* but that *the process ceases as soon as the vapor reaches a certain definite density* points to the conclusion that evaporation stops in the presence of a saturated vapor not because of its *pressure* but because *molecules passing from it into the liquid compensate for those escaping into the vapor.*

448. Boiling. When a liquid is exposed in an open vessel it evaporates more or less at all temperatures, since, as we have just seen, the pressure of air on its surface, even though it may be greater than the vapor pressure of the liquid, cannot prevent evaporation, though it retards it. This evaporation takes place more rapidly the greater the vapor pressure of the liquid, hence the greater volatility of ether than of water; so also water at high temperatures evaporates more rapidly than at low. If, however, the liquid be heated sufficiently bubbles of vapor will form in its

interior and rise to the surface and escape. The liquid is now said to be *boiling* or in a state of *ebullition*, and its temperature remains constant so long as it keeps boiling at a given pressure.

Boiling can take place only when the pressure of the vapor is equal to that of the atmosphere on the liquid surface, otherwise bubbles could not be formed.

The tabulated *boiling point* of a substance is that temperature at which its vapor pressure is equal to the standard atmospheric pressure, viz., 76 cms. of mercury.

If boiling takes place in a *closed* boiler the temperature does not remain constant but rises as the pressure of the contained air and vapor increases.

BOILING POINTS AT ONE ATMOSPHERE PRESSURE

Zinc	958.0° C.	Ammonia	−33.6° C.
Sulphur	444.5°	Carbon dioxide (sublimes)	−78.0°
Mercury	357.0°	Oxygen	−182.5°
Water	100.0°	Nitrogen	−195.0°
Alcohol	78.0°	Hydrogen	−252.0°
Ether	34.6°		

449. Effect of Pressure on the Boiling Point of Water. From the previous paragraph it appears that a table of vapor pressures shows the boiling points corresponding to different pressures. Referring to the table on page 307, it will be seen that near 100° C. the boiling point of water changes by one-tenth of a degree for 2.7 mm. change in pressure.

A vessel of water under the bell jar of an air pump may be made to boil by exhausting the air until the pressure·is slightly less than the vapor pressure of the water as given in the table.

On high mountains the temperature of boiling is so low that eggs cannot be cooked, and in some high altitudes closed vessels provided with safety valves are used in cooking in order that it may be possible to heat the water to a sufficiently high temperature. On Mt. Blanc water boils at

FIG. 235. Boiling at reduced pressure

84° C. By observing the boiling point the barometric pressure, and hence the height of a mountain, may be estimated. This process is known as hypsometry.

If a flask half full of water vigorously boiling is taken from the heating flame, and instantly corked air-tight with a rubber cork and inverted as shown in the figure, it may be made to continue boiling by cooling the upper part of the flask, for in this way vapor is condensed and the pressure on the interior is so much reduced that the liquid boils even though its temperature is decidedly below 100° C.

450. Geysers. The explanation of the action of geysers given by Bunsen is based on the dependence of the boiling point on pressure. Suppose a deep fissure or well into which water flows at the bottom and where it is gradually heated from below. The boiling point at the surface is 100°. At 36 cms., or a little more than 1 ft. below the surface, the added pressure of the water column will make the boiling temperature 101°. Suppose at this point the actual temperature is 100°. Again at 10 ft. below the surface the boiling point is about 107° and the actual temperature may be a little less; and at 50 ft. below the surface the boiling point will be 127° and the actual temperature may here also be supposed a little below this. Thus at each point the temperature of the water column filling the well will be just below that required to produce boiling. Now suppose that at some point down in the well the water becomes heated to its boiling point. The bubbles of steam in forming lift the whole upper column of water quickly so that each point in the column finds itself under a pressure less than that at which it boils and instantly steam bursts out at every point driving the mass of water violently out of the well.
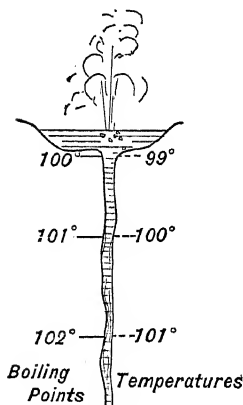
Fig. 236. Geyser

451. Effect of Dissolved Salts on Vapor Pressure. When a liquid contains salts in solution its vapor pressure at a given temperature will be less than in case of the pure solvent, the amount of the change depending on the nature and concentration of the solution. A saturated solution of common salt in water boils at 108.4° C., while one of calcium chloride boils at 179.5° and contains 76.4 per cent of the salt. The vapor escaping in these cases is pure water vapor at a pressure of 76 cms., and though the temperature of the bubbles of steam as they escape from the liquid may be higher

than 100°, it was discovered by Rudberg that *a thermometer in the steam a little above the liquid will record exactly as it would in steam from pure water.* The vapor as it escapes has a pressure of 76 cms., if that is the external pressure, and a temperature *above* 100° C.; it is, therefore, *non-saturated* and at once begins to cool, but when it reaches 100° it is saturated and any

further loss of heat causes condensation on the upper part of the vessel through which the steam is escaping, so that the vessel is soon heated to 100° and the escaping steam is kept at that temperature.

On the other hand if steam at 100° is continuously passed into a solution of salt, as in figure 237, the salt solution will be heated up to its own boiling point though that may be several degrees above 100° C.

452. Latent Heat of Vaporization. Heat is required for evaporation, just as for melting. The molecules of the liquid are torn
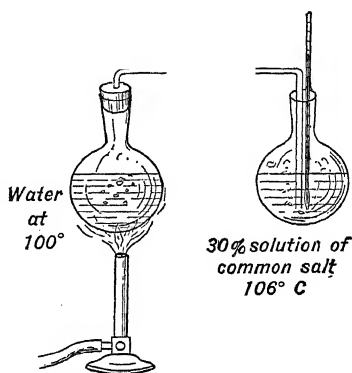
FIG. 237. Heating by condensed steam

away from each other in opposition to their cohesion and this requires work. Hence vapor has more energy than an equal mass of the liquid at the same temperature. *The heat required to change one gram of liquid into vapor at the same temperature is known as its latent heat of vaporization.* If dry steam is passed into a condensing vessel made of thin metal and surrounded by water in a calorimeter, as shown in figure 238, the steam will condense in the worm tube and collect in the bottom of the condenser. In condensing, its latent heat of vaporization is given up, and the condensed water is cooled from 100° to the final temperature of the calorimeter. If from the whole heat received by the calorimeter we substract the heat given out by the water in cooling *after* it is condensed, the remainder is the latent heat obtained from the condensation of the steam. The amount of condensed water is obtained by weighing the inner apparatus before and after the experiment, and so the heat per gram of condensed steam may be found.

FIG. 238. Calorimeter for vapor

The latent heat of vaporization of water is less at high temperatures than at low as might be expected; thus to evaporate 1 gram of water at 100° takes 536.6 gram-calories of heat, while 596.7 calories are required if the evaporation takes place at 0° C.

The latent heat L required to vaporize a gram of water at any temperature t may be determined from the following formula which expresses the results of Griffith's experiments,

$$L = 596.73 - 0.601t$$

where t is the temperature of evaporation on the Centigrade scale.

HEATS OF EVAPORATION (AT NORMAL BOILING POINTS)

Water	537	Ether	91
Wood alcohol	264	Ammonia	326
Alcohol	208	Liquid air	51

453. Joly's Steam Calorimeter. Dr. Joly has devised a very useful method of measuring specific heats which is based on the fact that if a body is immersed in steam at 100° condensation will take place on the body until its temperature is raised to 100° when no further condensation will take place. The heat received by the body is the latent heat of vaporization given out by the steam in condensing. The apparatus is shown in figure 239. The body the specific heat of which is to be determined is hung from the pan of a delicate balance, and the amount of water condensing on it is found by its gain in weight when steam is passed through the inner vessel.

FIG. 239. Steam calorimeter

Let w be the mass of condensed water, t the original temperature of the suspended body, m its mass, s its specific heat, and L the latent heat of vaporization of steam; then

$$ms(100° - t) = Lw.$$

454. Cooling by Evaporation. If a vessel of water is placed under the bell jar of an air pump and the air exhausted, the

water will boil as the pressure is reduced and at the same time its temperature will rapidly fall owing to its giving up heat energy to supply the latent heat of vaporization of the vapor coming off. The process may even be carried so far as to freeze the water. This is illustrated by a device due to Wollaston and known as a *cryophorus* (cold transferrer) which is shown in figure 240. It consists of a tube having a bulb at each end and containing only water and its vapor, the air having been driven out by boiling the water before sealing the tube. The water is all run into the upper bulb while the lower one is surrounded by a freezing mixture. Vapor arising from the water in the upper bulb takes away heat in evaporation and is condensed in the lower bulb. As this process continues the water in the upper bulb is soon frozen if protected from outside sources of heat by being surrounded with cotton.
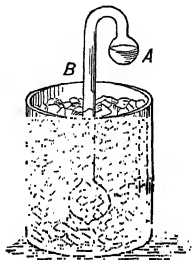
FIG. 240

In tropical countries water is kept in porous earthen jars put in a breezy place protected from the sun. The rapid evaporation from the moist surface of the jars keeps the water cold. If a few drops of ether or alcohol are poured on the hand the chilling as it evaporates is very noticeable. If a test tube partly filled with ether is placed in a glass of ice-cold water and the ether rapidly evaporated by causing a stream of air to bubble through it by a foot bellows, a thick shell of ice will soon be frozen around the test tube.

The ordinary wet bulb thermometer, § 457, affords another instance of cooling by evaporation.

455. Refrigerating Machines. The refrigerating machines used for ice making on a large scale and for cooling rooms for cold storage depend on the condensation and evaporation of ammonia, the arrangement employed being as follows. Ammonia gas is compressed by a pump into a condenser B where the heat developed by the compression and condensation is removed by a stream of water. From the condenser it is conducted in liquid form through a pipe leading to the region to be cooled. There it escapes through a valve C into a long pipe D which winds about the room to be cooled, and affords large

cooling surface. In this pipe the ammonia evaporates and expands and so takes up heat from the surrounding region which is accordingly chilled. The expanded gas is conducted back to the pump where it is again compressed into the condenser.
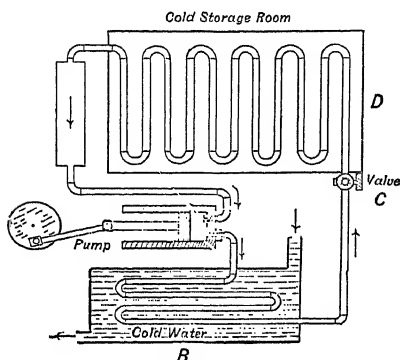
FIG. 241. Refrigerating by ammonia

The valve C has only a small opening and chokes the flow so that while the pressure in the pipe leading from the condenser to C is sufficient to keep the ammonia in the liquid form, beyond C the pressure is very small, permitting rapid evaporation and expansion.

456. Sublimation. A solid may evaporate directly without passing through the liquid state. This is known as *subliming*. Gum camphor sublimes very freely, also the naphthaline balls so often used for protection from moths. Ice slowly evaporates when below the freezing point, and carbon dioxide not only passes directly into vapor from the solid form but *cannot exist in the liquid state* at atmospheric pressure.

Heat is absorbed or becomes latent in sublimation just as in other changes of state, the latent heat of sublimation being the heat required to cause one gram of the substance to sublime at a given temperature.

457. Atmospheric Moisture. The determination of the moisture in the atmosphere is known as *hygrometry*, and is of much importance in meteorology.

When a mass of dry air in the free atmosphere receives water vapor the added pressure of the vapor causes the whole to expand, since its pressure cannot be greater than that of the surrounding atmosphere. The expanded air is less dense than dry air at the same pressure and temperature, for the water vapor which it contains is only $\frac{5}{8}$ as dense as the dry air which it displaces.

The following are some methods employed in determining the moisture in the atmosphere:

1. A measured volume of air is drawn through a tube containing some drying substance, such as calcium chloride or phosphoric anhydride, and the gain in weight of the drying substance gives the amount of moisture which the air contained.

2. A bright polished metal vessel is cooled till moisture from the air just begins to condense on its surface. The temperature at which this occurs is known as the *dew point*. If the dew point is found to be 10° then the pressure of water vapor in the air is such that it is saturated at 10°. Hence the vapor pressure is 9.1 mm. as shown in the table on page 307 which gives the pressure of saturated water vapor at different temperatures.

3. The temperature of the air as read by a wet bulb thermometer, one having the bulb covered with a thin piece of cloth kept wet by a wick dipping into a vessel of water, may be compared with the atmospheric temperature as given by a thermometer with a dry bulb.

The wet bulb thermometer will read lower than the one with dry bulb in consequence of evaporation, and the more rapid the evaporation the greater the difference in temperature will be.

When the two temperatures are known the hygrometric state of the atmosphere may be determined by reference to *psychrometric tables*.

The method is exceedingly convenient but is not reliable unless the air is drawn by the wet bulb thermometer at a regular rate, or unless the thermometer is whirled through the air with sufficient velocity to secure the maximum evaporation.

4. A human hair when treated with ether to remove oily substances is very sensitive to moisture, elongating when moist and contracting as it dries. When one end of such a hair is wrapped around a slender axle to which a pointer is attached, the varying moisture condition of the air may be read by the motion of the pointer; and the instrument is known as a hair hygrometer.

A piece of catgut when stretched by a light weight twists and untwists as the moisture in the air varies.

458. Humidity. The sensation of dampness is due to the degree of saturation of the air. When the air is cold a comparatively small amount of moisture will make it feel damp, because at a low temperature but little moisture is required to make a saturated vapor. Only 4.7 grams per cubic meter are required at the freezing point, while at 77° F. or 25° C. 22.75 grams may be contained in a cubic meter before saturation.

On this account the *relative humidity* is usually sought in meteorological observations.

The hygrometric state, or relative humidity of the air, is the ratio of the water vapor actually contained in a volume of air to the amount that it would contain at the observed temperature if the vapor were saturated.

PROBLEMS

1. If a Bunsen burner can heat 2 kgms. of water from 10° to 80° in 10 minutes, how much water can it boil away per hour?

2. How many grams of water are required to fill a room $3 \times 5 \times 4$ meters in size, with saturated water vapor at 20° C.? Water vapor has $\frac{5}{8}$ the density of dry air at the same temperature and pressure.

3. The barometric height on Mt. Washington is about 60.5 cms.; at what temperature will water boil there?

4. A flask half full of water boiling vigorously is corked tight and immediately immersed in a bath having the temperature 50° C. What will the pressure in the flask become, and when will it stop boiling?

5. When water boils at a pressure of 35.55 cms. of mercury, find the temperature and also the heat that must be supplied to evaporate 100 gms.

6. What is the weight of a cubic meter of saturated steam at 100° C. if water vapor has $\frac{5}{8}$ the density of dry air at the same temperature and pressure?

7. Find the temperature of the water at the bottom of a pail of water 30 cms. deep which is boiling while the barometer stands at 76.

8. Find the relative humidity when the air is at 20° C., the dew point having been found to be 10° C.

9. How much coal is needed to evaporate a cubic foot of water (28.3 kgm.) in a boiler at atmospheric pressure and boiling temperature, supposing the efficiency of the boiler to be 50 per cent and the heat of combustion of coal to be 8000 gram-calories per gm. of coal?

10. Find the total amount of heat required to change 100 gms. of ice at −20° C. into steam at 100°.

11. A mass of 100 gms. of copper at 20° is suddenly enveloped in steam at 100°. Find the amount of steam that will condense on the copper.

12. A preserve jar containing only water and its vapor is sealed up and put into a kettle of water which is kept boiling. Will the jar burst, and what will the pressure within it become?

13. A preserve jar half full of water and half full of dry air at pressure 76 is sealed up at temperature 20° C. and put into a kettle of water which is kept boiling. Find what the pressure in the jar will become, neglecting the expansion of water and glass.

14. In the cryophorus how much water must distil over from the upper to the lower bulb in order that 50 gms. of ice may be frozen? And what value of the latent heat of vaporization should be used?

15. How much ammonia must be evaporated per hour in the refrigerating coils in a cold-storage room if the temperature is to be maintained at 0° C. when 500 gram-calories of heat are flowing into the room per second? *

* Heat of evaporation of ammonia = 301.8 gram-calories at 0° C.

16. How much heat in British thermal units is required to evaporate a pound of water at 100° C.?

17. How much coal is required to evaporate 100 kgms. of water in a boiler in which the gauge pressure is maintained at 54.4 lbs., supposing no heat wasted?

CONDENSATION OF GASES

459. Triple Point. The particular temperature and pressure at which a substance may exist as solid, liquid or vapor, is called the *triple point*. If a vessel containing only water and its vapor is cooled until the water begins to freeze it will then be at its triple point, for all three states or phases (solid, liquid, and vapor) are in equilibrium with each other.

In the pressure-temperature diagram for water (Fig. 242) the curve of *boiling points* between the liquid and vapor regions shows the condition of temperature and pressure at which the vapor is *saturated* and in equilibrium with its liquid.
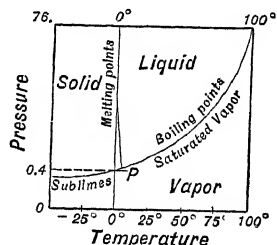
FIG. 242. Triple point

The curve between the solid and liquid regions shows melting points at different pressures, while that between the solid and vapor regions shows the pressure of the vapor in equilibrium with ice at different temperatures.

The triple point is where these lines meet. In case of water the melting point rises very slightly as pressure is diminished, so that at the triple point the temperature is about 0.0075° C. above zero, and the pressure is 4.6 mm. of mercury.

When the temperature and pressure of a substance are such that the point representing this state in a temperature-pressure diagram, similar to that of figure 243 for water, does not lie on any of the curves then the substance must all exist either as a solid, liquid or vapor; depending upon which of the three areas in the figure contains the point.

In case of carbon dioxide the pressure at the triple point is greater than one atmosphere, hence that substance cannot exist in the liquid form at atmospheric pressure.

460. Condensation of Carbon Dioxide and Critical Point.
The transition from the gaseous to the liquid state was first
thoroughly studied at different temperatures and pressures by
Andrews (1863), a diagram of whose results is given in the figure.
Each point on the diagram corresponds to a certain state of the
substance. The abscissa or
distance of the point from
the side line gives the
volume in units of the scale
at the bottom, while the
distance of the point from
a horizontal base line meas-
ured by the scale at the side
gives the pressure in at-
mospheres. The base line
corresponding to *zero* pres-
sure is far below the dia-
gram, which is made only
large enough to include the
actual observations. A
line on the diagram repre-
senting the series of states
through which the sub-
stance may be put at a given
temperature is called an
isothermal line. Thus the
line marked 13.1° indicates
that if a gram of CO_2 be
taken at 13.1° and at a
pressure less than 50 atmospheres the volume will be greater than
7.1 c.c. As the volume is diminished the pressure increases until
when the volume is 7.1 c.c. the pressure becomes 50 atmospheres,
but at this point condensation begins and the pressure remains
constant until all is condensed. This part of the isothermal line,
where the substance is part liquid and part saturated vapor, is
horizontal, since the pressure is constant.
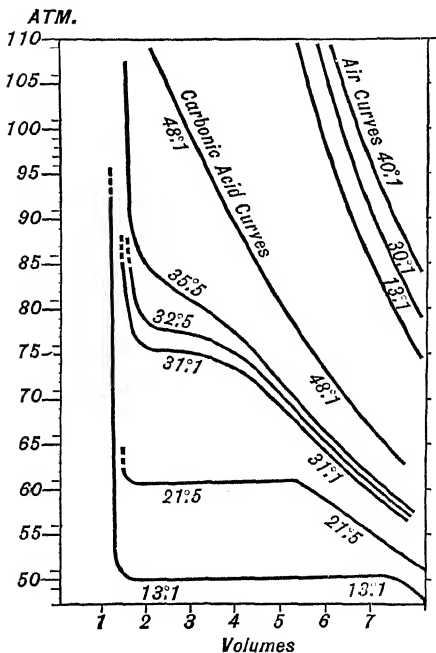
After the vapor is entirely condensed any further decrease in
volume is accompanied by rapid rise in pressure as indicated by
the nearly vertical branch of the isothermal line. Where the

FIG. 243.　Isotherms of CO_2

isothermal line is horizontal the substance is part liquid and part saturated vapor, becoming wholly saturated vapor at the right end where the line begins to drop from the horizontal; beyond this point the vapor is non-saturated, departing considerably from Boyle's law at first, but conforming to it more closely as its volume increases and pressure diminishes.

At 21.5° it is noticeable that the volume of the saturated vapor is less, about 5.2 c.c., while the volume of the liquid is greater than at the lower temperature, and condensation does not occur till the pressure has reached 60 atmospheres, this being its vapor pressure at 21.5° C.

Thus as the temperature is raised the density of the saturated vapor increases while that of the liquid decreases until at 31° they seem to come together, the saturated vapor having the same density as the liquid. In this case there is no visible condensation with separation of liquid and vapor, and the isothermal line shows no straight horizontal part, but simply a point of inflection where the tangent is horizontal.

The next isothermal, that for 35.5°, shows a point of inflection, a point where the compressibility is a maximum, as indicated by the large decrease in volume for small rises in pressure, but there is no condensation. And at 48° there is scarcely any evidence even of a point of special compressibility, the pressure rising steadily and rapidly as the volume is diminished.

The point at which the density of the liquid becomes equal to that of its saturated vapor is called the critical point. The temperature and pressure of the substance at that point are known as its *critical temperature and pressure,* and the volume of one gram as its *critical volume.*

The critical temperature may also be defined as *that temperature above which the substance cannot exist as a liquid having a free surface.*

461. Gas and Vapor. It thus appears that there is no sharp distinction between gases and vapors. What are ordinarily known as gases are substances whose critical temperatures are so low or critical pressures so great that under ordinary conditions they cannot exist as liquids with a free surface, while vapors arise from substances whose critical temperatures are so high that they ordinarily exist even at atmospheric pressure in the liquid state.

CRITICAL TEMPERATURES, PRESSURES, AND VOLUMES

SUBSTANCE	TEMPERATURE IN DEGREES C.	PRESSURE IN ATMOSPHERES	VOLUME OF 1 GRM. IN C.C.
Water.....................	365.0°	195.0	2.3
Ether.....................	194.4	35.6	3.8
Sulphur dioxide............	155.4	78.9	1.9
Ammonia..................	130.0	115.0	...
Carbon dioxide............	30.92	77.0	3.4
Oxygen...................	−118.0	50.0	1.5
Nitrogen..................	−146.0	35.0	2.7
Hydrogen.................	−242.0	20.0	...
Helium...................	−266.0	2.3	...

462. Condensation of Gases. Faraday, about 1823, began a series of experiments in which he liquefied nearly all the known gases except oxygen, nitrogen, hydrogen, and carbon monoxide.
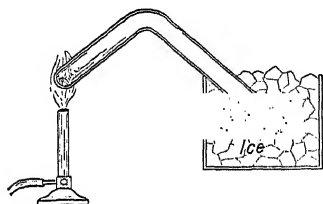
FIG. 244

The form of apparatus used by him for chlorine and other gases is shown in figure 244. It consists of a strong bent glass tube hermetically sealed, in one end of which is placed the substance or mixture from which the gas is to be evolved, while the other end is placed in a freezing mixture to induce condensation. When heat is applied the gas is given off on one side and the pressure due to its own evolution causes it to condense on the other. In cases where heat was not required the two ingredients were placed separately in the two branches of the tube which was then sealed. The tube was then tipped up so that the substances were mixed in one branch of the tube while the other was introduced into the freezing mixture as before.

Solid Carbon Dioxide. When carbon dioxide, after being liquefied by pressure, is cooled, and then allowed to escape from a small opening, the evaporation and expansion produce such a degree of cold that a considerable part of the escaping substance is frozen into snow. If the jet is enclosed in a woolen bag this snow may be collected. It slowly sublimes, passing directly from the condition of solid to vapor only so fast as the necessary heat of vaporization is obtained from surrounding bodies.

The temperature of solid CO_2 at atmospheric pressure is $-78°$ C.; if a little is placed on the hand it is kept from close contact at first by the gas given off due to the heat of the hand. If pressed into contact it burns like a hot iron. If mixed with ether a freezing mixture is obtained giving a temperature of $-78°$ C. at atmospheric pressure, and if the pressure is reduced by an air pump $-116°$ C. may be reached. The mixture even at atmospheric pressure readily freezes mercury.

463. Liquefaction of Air. Many attempts were made to liquefy the permanent gases, as they were called, by Faraday, Natterer, and others, but without success until, in 1878, Cailletet and Pictet, working independently, one at Paris and the other at Geneva, almost simultaneously achieved the desired result. Both subjected the gases to great pressure and then cooled them to the lowest point attainable by evaporating liquid sulphur dioxide or carbon dioxide under diminished pressure. But even at the low temperatures thus secured no condensation was observed until a stopcock was opened and the compressed gas suddenly permitted to expand. The cooling due to this sudden expansion caused a cloud of particles of condensed gas to appear. In this way oxygen, nitrogen, and carbon monoxide were shown to be liquefied.

WROBLEWSKI and OLSZEWSKI obtained a still lower temperature by cooling liquid ethylene first with ice and salt, then with carbon-dioxide snow mixed with ether, and finally the cooled ethylene contained in a triple-walled glass vessel was made to boil at diminished pressure, the gas as it evaporated being pumped out by an exhaust pump. (Fig. 245.) In this way a temperature

FIG. 245. Wroblewski's apparatus

of $-136°$ C. was reached, and when oxygen was compressed into a tube dipping below the surface of the boiling ethylene it was condensed at a pressure of about 20 atmospheres. In this manner considerable quantities of liquid oxygen and nitrogen were first obtained.

Linde's Apparatus. The present methods of obtaining liquid air on a large scale are based on the progressive cooling of a stream of escaping gas by its own expansion, and the first apparatus of this kind was devised by Dr. Linde in 1895. Compressed air at a pressure of about 200 atmospheres, and dried and purified from
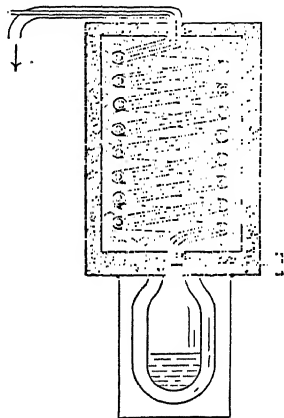
FIG. 246. Liquid air interchanger

carbon dioxide, passes into the inner tube of the *interchanger* which contains long coils of tubing one within the other and packed in felt to prevent the inflow of heat from outside. At the lower end of the interchanger is a needle valve through which the compressed gas is allowed to escape in a steady stream. The escaping gas cooled by expansion passes out through the outer tube of the interchanger, thus cooling the inflowing stream of compressed gas in the inner tube. But this on expansion is still further cooled and so the cooling goes on progressively until the temperature becomes so low that a part of the air

is liquefied as it escapes at the valve and falls into the receiver below, where it is collected. (See note, § 465.)

This receiver is a double-walled glass vessel (such as are used in thermos bottles) known as a Dewar flask. The space between the walls of the flask is thoroughly exhausted of air to prevent conduction of heat to the inner vessel. The surface of the inner vessel is coated with a silver film which forms a bright metallic mirror that reflects radiation and thus still further aids in preventing the liquid air from receiving heat from outside.

Liquid air when first produced contains both oxygen and nitrogen, but as the boiling point of nitrogen ($-195.5°$ C.) is lower than that of oxygen ($-183°$ C.) the former soon boils off leaving nearly pure oxygen.

464. Liquefaction of Hydrogen and Helium. The liquefaction of hydrogen has been accomplished by Dewar using an improved form of Linde's apparatus in which two separate interchangers were used, one entirely surrounded by the other. The

outer one was first used for the production of liquid air and in this way the whole apparatus was cooled to $-180°$ C. Hydrogen was then passed through the inner apparatus and still further cooled by its own expansion until it finally collected as a clear liquid boiling under atmospheric pressure at $-252°$ C. or 21° above the absolute zero.

On reducing the pressure the temperature of the boiling hydrogen was lowered until it froze into a solid at $-258°$ C.

A cubic centimeter of liquid hydrogen weighs 0.086 gm.; it is therefore the lightest liquid known.

If a bulb containing air has a long neck which is sealed up and surrounded by liquid hydrogen the air will condense and freeze in the neck leaving the bulb highly exhausted.

If fragments of box charcoal are contained in the cooled neck their absorption is so powerful that the bulb becomes almost a perfect vacuum.

Helium, the last gas to yield to condensation, was finally liquefied in 1908 by the Dutch physicist, Onnes. Liquid helium, according to Onnes, boils at $-268.5°$ C., and has a density 0.15.

465. Note on Cooling by Expansion in Linde's Apparatus. The cooling of a steady stream of gas escaping under pressure through a small opening, as in Linde's apparatus for the liquefaction of air, is by no means as great as when a mass of gas is expanded in a non-conducting cylinder as explained in § 423. For while the expansion of the gas tends to cool it, the kinetic energy of the gas rushing out of the opening tends to heat the expanded gas and one effect nearly balances the other so that the cooling is but slight in case of air at room temperature. As air is cooled, however, the effect increases until a sufficiently low temperature is reached to liquefy the air.

When a stream of *hydrogen* gas at room temperature is forced in this way through a small opening the gas is slightly *heated* instead of being cooled and it is only after being cooled below $-80°$ C. to begin with, that the escaping jet is cooled at all by its own expansion.

MELTING AND BOILING POINTS OF CONDENSED GASES

SUBSTANCE	MELTING POINT	BOILING POINT
Helium........................	Unknown	$-268.5°$ C.
Hydrogen......................	$-258°$ C.	-252.0
Nitrogen......................	-210	-195.5
Oxygen.......................	-227	-183.0
Air...........................	....	-191.0

Heat Engines

466. Heat Engines. The conversion of heat into mechanical energy is of the greatest importance to man, since vast stores of fuel existing in the earth as coal, petroleum and gas are thus made available for useful work.

It is interesting to consider that these deposits are really storehouses of the energy of sunlight which fell on the earth in ages long gone by and effected the separation of carbon from oxygen in plants, thus storing up potential energy which is ready to be given back to us as energy of heat under the magic touch of flame.

The principal kinds of heat engines are the steam engine, the hot-air engine, and engines that burn gas or vapors explosively.

467. The Steam Engine. A simple double-acting steam engine is shown in the diagram (Fig. 247). Steam from the boiler is admitted to the steam chest S, passes through one steam port into the cylinder A and forces the piston toward the left. Whatever steam or air is in the other end of the cylinder B escapes through the other port, passes under the cup-shaped slide valve and out at the exhaust E. But the slide valve is so connected to the main shaft through the eccentric that as the piston moves toward the left the slide valve moves toward the right, closing the first port and opening that at the other end of the cylinder. Steam is thus admitted first into one end of the cylinder and then into the other, forcing the piston back and forth. The flywheel F which has a large moment of inertia steadies the motion and carries the crank C past the "dead centers."

468. High-pressure and Condensing Engines. When the exhaust E opens directly into the atmosphere the engine is called *high pressure*, because the power depends on the excess of the steam pressure in the boiler above the atmospheric pressure outside. Ordinary locomotives and most small engines are of this type.

But greater economy is obtained by connecting the exhaust to a vacuum chamber in which the steam as it comes from the engine is condensed by a jet of cold water or in tubes surrounded by cold water, a small pump being provided to pump out from the vacuum chamber the condensed steam as well as any air that may

have leaked in. Such engines are known as *condensing* engines, and since the pressure in the vacuum chamber may be less than 1 lb. to the square inch the back pressure against the piston is less by 14 lbs. to the square inch than if the exhaust had opened into the atmosphere, and the effective pressure is consequently just so much greater.

469. Compound Engines. To get as much work as possible out of steam it should be used *expansively* in the engine, and
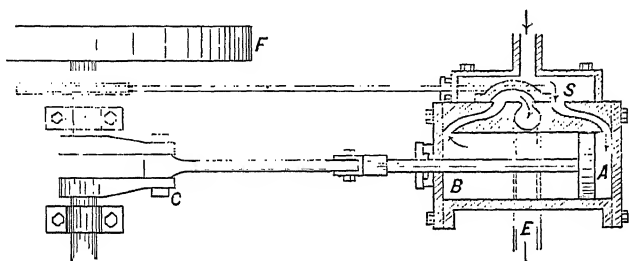
FIG. 247. Steam engine

should not pass into the exhaust until its pressure in consequence of expansion has diminished almost to that in the exhaust, otherwise it escapes with explosive puffs which represent lost energy.

If the expansion takes place in one cylinder the steam which is admitted at high pressure and temperature does not escape until its pressure and temperature are both greatly reduced by expansion. *To avoid this great change in pressure and temperature in a single cylinder, compound engines are used in which the steam passes successively through several cylinders, a part of the expansion taking place in each one.* Each cylinder must be larger than the preceding one to allow for the expansion of the steam, and as the steam pressure in one cylinder is less than in the preceding one the area of the piston head is made correspondingly larger in the second, so that the total force exerted by each piston may be about the same.

470. Steam Turbine. In the DeLaval steam turbine one or more jets of escaping steam are directed against a series of blades set in the rim of a wheel, driving it with great velocity.

In the Parsons turbine (Fig. 248) the blades are set in rows or bands around the circumference of a long cylindrical drum, which rotates inside of an outer case. Steam is admitted around one end of the cylinder and impinges obliquely on the first row of
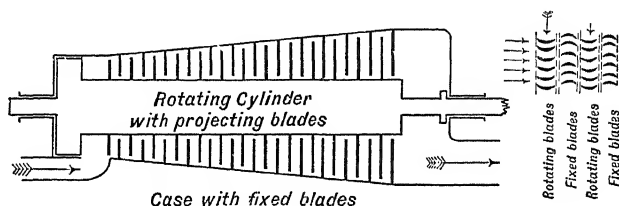
Fig. 248.   Steam turbine

blades. These blades are curved so that they deflect the stream of escaping steam as it passes between them and direct it against a second row of blades fixed to the outer case. These in turn deflect the stream so that it strikes obliquely against the second row of blades on the rotating cylinder, thus the escaping steam acts on row after row of blades successively from one end of the cylinder to the other where it escapes into the exhaust. The space between the rotating cylinder and the outer case widens from one end toward the other to allow for the expansion of the steam as it passes through, and consequently the blades are short at the end where the steam enters but are longer as the other end is approached. The turbine is thus in some respects like a compound engine, the successive rows of blades in the former corresponding to the successive cylinders in the latter.

1. Drawing in explosive mixture

2. Compressing the mixture

3. Piston driven out by explosion

4. Pushing out waste gases
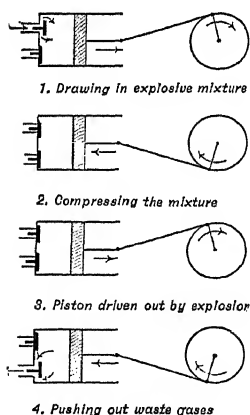
Fig. 249.   Four-cycle gas engine

471. Gas and Gasolene Engines. In these engines an explosive mixture of gas or gasolene vapor and air is drawn into the cylinder and there ignited, power being obtained from the expansive force of the hot gases which result from the explosion. Figure 249 shows the series of operations in the "four cycle" or Otto type of engine. In the first

outward stroke the mixture of air and gas in proper proportion is drawn in; the valve then closes and the mixture is compressed on the return stroke. When the crank is on the " dead center " and the compression is maximum the mixture is ignited by flame or electric spark and power is obtained from the thrust of the expanding gas on the outward stroke. The exhaust valve then opens and the waste gases are driven out as the piston moves back. The engine is made *single acting* to avoid undue heating and the cylinder is also kept cooled by a circulation of water around it. It will be observed that power is obtained only in the third operation, or on every alternate outward stroke. A heavy flywheel is, therefore, used, or in automobiles four such engines may act on one shaft, the explosions taking place successively in the several cylinders, one to every half revolution of the shaft.
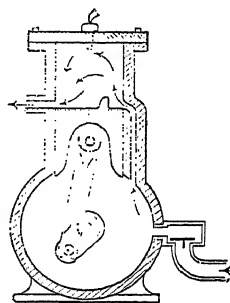
FIG. 250. Two-cycle gasolene engine

In so-called " two-cycle " engines the explosive mixture which has been compressed in the crank case by the outward movement of the piston is admitted to the cylinder near the end of the stroke while the exhaust valve is open, and by its inrush helps to displace and drive out the spent gases. The mixture is compressed as the piston moves back and then exploded, giving power on the outward stroke. Of course in this case there is some mixture of the spent gases with the fresh charge and some of the latter is also lost through the exhaust.

472. Efficiency of Heat Engines. The efficiency of an engine is the ratio of the work done in a given time to the mechanical equivalent of the heat which is supplied to it during that same time. It is shown by thermodynamic reasoning that no engine can have a greater efficiency than

$$T - T_0$$

where T is the highest temperature of the working substance as it passes through the engine and T_0 is its lowest temperature, both measured on the *absolute scale.*

Thus if an engine takes in steam at 163 lbs. pressure, the temperature of which is 185° C. and if the temperature of the exhaust is 100° C., then $T = 458$ and $T_0 = 373$ and its efficiency cannot be greater than $\frac{85}{458} = 18\% +$.

Taking account also of the loss of heat in the furnace and boiler, it is found that a good engine (multiple expansion and condensing) may give about 1 horse-power-hour per pound of coal, but ordinary non-condensing engines require 2 or 3 lbs. of coal per horse-power-hour.

For a further discussion of the relations of heat to work, taking up the Second Law of Thermodynamics, Carnot's cycle, and the Kelvin absolute or thermodynamic scale of temperature, see *Appendix I.*

REFERENCE

See account of development of the steam engine in *Heat as a Form of Energy*, by R. H. THURSTON.

RADIATION AND ABSORPTION

473. Radiation. A person standing near an open fire is conscious not only of the light coming from it, but also of a sensation of warmth which is felt in the skin wherever it is directly exposed to the glow. This sensation is lost when an opaque screen is interposed, and returns as instantaneously as the light when the screen is withdrawn. This is shown by the fact that after a solar eclipse the warming effect of the radiation from the sun reappears as soon as the light itself.

The radiation may be felt even through a sheet of thin ice and by means of a lens of ice it may be converged into a focus sufficiently intense to ignite gun cotton.

But since both sides of the ice are at the same temperature (the temperature of melting ice) no heat can be transmitted by ordinary conduction. Radiant energy is, therefore, transmitted by a very different process. This is also shown by the fact that it passes with the greatest facility through a vacuum.

The process of emitting energy in this way is called *radiation*, and the total stream of energy coming from the body in this way is called its radiation.

We shall discuss here some circumstances which influence the

giving out and absorbing of this radiant energy, but radiation itself, its nature and varied phenomena, will be taken up in our later study of light, for light is but that part of the total stream of radiation to which the human eye responds.

474. Instruments for Detecting Radiation. The heating effect of radiation is detected usually either by the thermopile, radio-micrometer, bolometer, or radiometer. The thermopile and radio-micrometer are explained (§§ 675, 682) in the section on thermo-electricity.

In the bolometer, devised by Langley, a thin strip of platinum perhaps 0.01 mm. thick and 0.5 mm. wide and having a blackened surface, is mounted in connection with a Wheatstone's bridge and galvanometer so that its resistance may be balanced. When radiation falls on the strip it is heated, and in consequence its electrical resistance changes slightly, which disturbs the balance of the bridge and causes a current to flow through the galvanometer. The mass of the platinum strip is so small that its change in temperature takes place almost instantaneously when radiation falls upon it. Langley was able to make the arrangement so sensitive that a change in temperature of the strip as small as one-millionth of a degree could be detected.

It was found by E. F. Nichols that the principle of the radiometer (§ 477) might be used in the construction of an instrument which was exceedingly sensitive for the detection and measurement of radiation.

In this instrument a light cross arm of wire carrying on each end a small disc of mica blackened on one side, and having a small mirror hung from it, is suspended by a fine quartz fiber in a vessel from which the air can be completely exhausted. The mica discs are vertical with their edges toward the axis of suspension, and the blackened sides of both face toward the same side. When a moderate exhaustion is reached the slightest radiation falling on the blackened side of one of the discs causes it to be repelled as explained in § 477. The suspended system consequently turns through a small angle which may be determined by observing through a telescope the image of a scale reflected in the mirror.

475. Radiating Power. Bodies at the same temperature may differ greatly in radiating power. This is shown by Leslie's cube, which is a brass cubical vessel containing boiling water.

This view is known as *the theory of exchanges.*

479. Apparent Radiation of Cold. The radiation from a block of ice may be converged upon a thermopile by means of a concave mirror and produces a decided cooling effect.

But whatever radiation goes from the ice to the thermopile must have energy and hence when absorbed must give *heat to* the thermopile. The explanation of the cooling is found in the theory of exchanges; for while the thermopile is receiving radiation from the ice, it is itself giving out more energetic radiation and is therefore cooled. The ice gives its feeble radiation to the thermopile but it intercepts the more intense radiation that would have reached the thermopile from other warmer bodies if the ice had not been there.

480. Equality of Radiating and Absorbing Powers. *The Stewart-Kirchhoff Law.* Imagine a body *A* supported at the center of a hollow vessel from which the air has been completely exhausted, and the interior surface of which is coated with lamp-black. If the outer vessel is kept at a constant temperature the inner body will also finally come to that temperature. It will then be in a state of equilibrium, *giving out just as much energy in radiation as it absorbs from the radiation that falls upon it.*

If the central body is a good reflector, like a piece of polished metal, it will reflect most of the radiation that falls upon it, absorbing only a small fraction. But its own radiation must exactly make up for what it absorbs, consequently it will radiate but little, and the *total* radiation coming from the body, being made up of what is reflected together with what the body radiates, must be just equal to the total radiation falling upon it.

If, on the other hand, the central body is a good absorber, such as a fragment of carbon, it will absorb nearly all the radiation falling upon it, reflecting very little. In this case it will radiate strongly, the radiation being equal to what it absorbs, and here also the total stream of radiation coming from the body is exactly equal to that which falls upon it, for its own radiation supplies the place of what it absorbs.

In a closed region, then, which is all at one temperature, the total radiation coming from any surface, partly reflected and partly radiated, is the same whatever may be the nature of the surface, whether it is a good reflector or a poor one, and is

equal to the radiation which would be given off at that temperature from a *perfectly absorbing body* or an *ideal black body*.

The above conclusion was reached independently by Kirchhoff and Balfour Stewart about 1858, it is illustrated by an experiment performed by Draper in 1847, in which a gun barrel containing fragments of various metals, colored crockery, etc., was heated in a furnace to a red heat. On looking into the gun barrel one substance could not be distinguished from another, the stream of radiation being the same from all, and equal to black-body radiation. Draper drew from his experiment the erroneous conclusion that all bodies became self-luminous at the same temperature, about 525° C.

The radiation which is actually emitted by a body as distinct from what is reflected by it is, therefore, equal to the portion of black body radiation that it can absorb at the same temperature.

Consequently bodies which are good reflectors and poor absorbers are also poor radiators, while poor reflectors which absorb strongly are also good radiators.

481. Law of Total Radiation. From the study of a large number of experimental results Stefan in 1879 concluded that the total energy of radiation R coming from a body was proportional to the fourth power of the *absolute* temperature, or

$$R = CT^4$$

where C is a constant.

Boltzmann, in a masterly discussion based on Maxwell's electromagnetic theory of light, reached the conclusion that Stefan's law is strictly true for *black-body radiation;* and this conclusion is borne out by the very thorough experimental investigations of Lummer and others. The law is also found to be approximately correct in other cases of purely thermal radiation.

The radiation in gram-calories per second from 1 sq. cm. of surface of a *black body* at temperature T reckoned from the absolute zero, is found to be $(1.30 \times 10^{-12})T^4$.

482. Law of Cooling. If a body at a temperature t is placed in an enclosure at some lower temperature t', it will cool, and the rate of cooling will be rapid if $t - t'$ is large. It was assumed by Newton that in such a case the rate of cooling is proportional to $t - t'$, or the heat lost per second equals $K(t - t')$, where K

is a constant to be determined by experiment. This law is nearly true if the difference between the two temperatures is not large, and it is often convenient to use. But the true law of cooling is based on the law of exchanges. By the preceding paragraph the heat given out in radiation by a body whose absolute temperature is T is equal to CT^4 where C is a constant. If
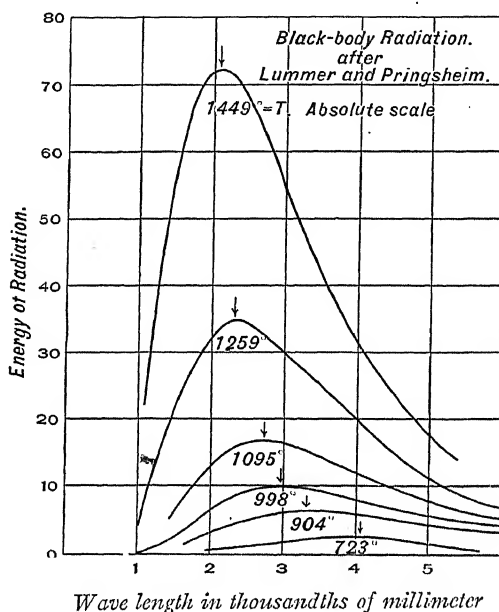
FIG. 253. Curves showing that the wave length of most energetic radiation is shorter in proportion as the temperature of the radiating body is higher

it receives radiation from a black body at temperature T_1 it will absorb CT_1^4, consequently the loss of heat per second is equal to

$$C(T^4 - T_1^4).$$

483. Wave Length of Most Energetic Radiation — Wien's Displacement Law. The radiation from a hot body, as we shall see later (§ 934), is complex in its nature and made up of ether waves of different wave lengths.

In figure 253 are given curves each of which corresponds to

a certain temperature and shows how the intensity of the radiation of a *black body* at that temperature varies with the wave length.

It will be observed that the hotter the radiating body the shorter is the wave length of most energetic radiation (corresponding to the highest points on the curves). Experiment and theory have combined to establish the remarkable law that the wave length of most energetic radiation is inversely proportional to the absolute temperature, or in symbols:

λT = a constant, found to be 2940 by Lummer and Pringsheim, where λ is the wave length in thousandths of a millimeter for the highest point on the energy curve and T is the corresponding temperature measured from the absolute zero.

Assuming that the radiation from the sun is sufficiently like that from a black body for the law to apply, we may determine its temperature. For the wave length of maximum energy in the sun's radiation is found by Langley to be 0.0005 mm., which gives for the sun's temperature $\dfrac{2940}{0.5}$ = 5880° absolute, or 5607° C.

Since the longest radiation waves that affect the eye give the sensation of red, the above law shows why a heated body should first become *red* hot and as the temperature rises the shorter waves become relatively more energetic until finally it appears *white* hot.

484. Quantum Theory. In order to account for the manner in which the energy of the radiation from a black body is distributed among the different wave lengths (§ 483) as shown by such curves as those in figure 253, Planck found it necessary to assume that there is something in the process of radiation which causes energy to be radiated in small units or packets called *quanta*, and that the elementary unit or quantum for any given wave length is equal to $h\nu$ where ν is the frequency of the emitted radiation and where h is an absolute constant known as *Planck's constant.* ($\nu = \dfrac{V}{\lambda}$ where V is the velocity of light and λ is the wave length of the radiation.) The value of h in C. G. S. units is found to be 6.56 × 10^{-27} erg seconds.

Since the time that Planck first proposed the quantum theory in 1901, many branches of experimental research on the emis-

sion and absorption of radiation have shown most convincingly that radiation is emitted and absorbed in minute grains of energy of size $h\nu$, called quanta. In fact the modern theories of the structure of the atom, as proposed by Bohr and others, are fundamentally based on the idea of quanta. The size of these quanta is seen to depend upon the wave length of the radiation. For long wave lengths, such as those of radio (§ 818) or of radiant heat (§ 483) the quanta are thousands to millions of times smaller than for the shorter wave lengths of light (§ 975) or the extraordinarily short ones of X-rays (§ 977).

Although called by different names, these waves are all of the same nature. They differ only in their wave lengths, power of penetration, etc.

The quantum theory and radiation have assumed such importance in modern physics that a special discussion of this subject is given at the end of the book.

485. Dew. Leaves and grass are rather good radiators and on clear nights they radiate strongly toward the sky and receive very little radiation in return. What is received comes for the most part from the air, which like all gases is a very poor radiator. Consequently vegetation is cooled and if there is much moisture in the air it condenses in the form of dew. Cloudy nights are unfavorable for the formation of dew since clouds radiate toward the earth.

When the temperature of the air is near the freezing point the chilling due to radiation causes ice crystals to form and *frost* is deposited instead of dew.

On windy nights there is usually no dew or frost because the rapid movement of the air over leaves and grass acts by *conduction* to keep vegetation at the same temperature as the general mass of air, thus the heat lost in radiation is supplied by convection and conduction; but on still nights the layer of air resting next to a cooled leaf soon becomes chilled below the average air temperature.

MAGNETISM

PROPERTIES OF MAGNETS

486. Natural Magnets. It was known to the ancients that certain iron ores had the power of attracting iron filings and small fragments of the same ore. The first specimens of this ore were obtained at Magnesia in Asia Minor and were on that account known as *magnets*. The mineral exhibiting this quality in the highest degree is a compound oxide of iron now known as *magnetite*. If such a natural magnet or *lodestone* is dipped into a mass of iron filings they cling to it in tufts especially at certain points called poles.

487. Mariner's Compass. If a lodestone having a strong pole at each end is balanced on a point or suspended by a cord or placed upon a float in water, it will set itself with one pole toward the north and one toward the south. The mariner's compass, which makes use of this property of the lodestone, was known in Europe in the year 1200 and probably earlier among the Chinese.

488. Artificial Magnets. If a small strip of hardened steel is brought into contact with a lodestone it becomes a magnet, and retains the property even when taken away. Iron filings will cling to it in tufts usually at its ends. If it is balanced on a point, one end will turn toward the north just as in the case of the lodestone. Such a piece of steel is said to be *magnetized* and to *exhibit magnetism*. When balanced on a point so that it can freely turn it is called a magnetic needle.

Very powerful magnets are made by causing a current of electricity to flow around a core of soft iron; such *electromagnets*, as they are called, will be discussed later (§ 693).

489. Magnets Have Two Kinds of Poles. *The fact that a magnetic needle will always set itself with the same pole pointing to the north indicates that the two poles are different.* If two magnetic needles are brought near each other it will be found

that the two north seeking poles repel each other; so also the two poles that turn toward the south repel each other; but if the north pole of one is brought near the south pole of the other decided attraction is observed. *Thus like poles repel and unlike attract each other.*

The pole turning toward the north is usually called the north pole in English books, but the French call it the south pole because its polarity must be like that of the south pole of the earth, considering the earth as a magnet.

490. Number of Poles. If a thin strip of hardened steel, a piece of clock spring for example, be magnetized by drawing a pole of a lodestone or other magnet over it from one end to the other it will probably be found to have two well-marked poles, one at each end. If we break the magnet in the middle and try to isolate one pole, it will be found that poles have appeared where it was broken and that each fragment has two opposite poles. *However small the magnet may be broken up, each piece shows a north pole and a south pole. No one has ever made a magnet with one pole.*

It is possible, however, for a magnet to have any other number of poles and a ring may be magnetized and have no poles at all. *Long thin bars of steel when magnetized often show more than two poles.*

491. Relative Strength of the Poles. When a magnetic needle is floated in a dish of water it at once sets itself in a north and south direction, but *it shows no tendency to be drawn toward the north or toward the south.* The north pole of the magnet is urged toward the north with a force equal and opposite to that acting on its south pole. The force between the earth and the north pole of the magnet is equal and opposite to that between the earth and the south pole of the magnet. It is concluded, then, that *the two poles of a magnet are equally strong.*

If the floating magnet has more than two poles, the result is the same, it is not drawn either toward the north or south. This indicates that *the combined strength of the north poles in a magnet is equal to that of its south poles.*

492. Nature of Magnetism. The fact that the fragments of a magnet always have two poles indicates that *magnetism is a condition which prevails throughout the whole mass of the mag-*

net, and polarity is merely an external manifestation of that condition. A piece of steel or iron is conceived as made up of particles or molecules each one of which is a little magnet. When the steel is not magnetized these particles are thought of as turned under the influence of their mutual attractions so as to form little closed groups in which the north pole of one particle is drawn toward the south pole of a neighboring one. At any point of the surface north poles and south poles thus neutralize each other so far as any external effect is concerned. Such a bar shows no evidence of poles, and we say it is not magnetized.

If, however, the bar of steel is placed between the poles of a powerful magnet, or if the opposite poles of two magnets are placed near together on the middle of the bar and then drawn apart toward its two ends, the particles of the bar are rearranged, being drawn apart from their former association under the

FIG. 254.   Non-magnetized steel bar   FIG. 255.   Magnetized bar

influence of the more powerful external attraction. The arrangement is now that shown in figure 255 where the particles are arranged in continuous chains or filaments, running from one end of the magnet toward the other. All the little south poles now have one general direction and all the ends of the filaments terminate on the end and sides of the magnet toward the right, while the north ends (point of the arrows) all terminate toward the left. The bar now exhibits north polarity on the left and south polarity on the right. If all the magnetic filaments were *straight parallel* chains of particles extending from one end to the other then polarity would be found only on the extreme ends, but in consequence of the mutual repulsion of similar poles the arrangement becomes as above shown, and the poles always extend over the sides of the magnet.

For a discussion of the part which electrons play in magnetization, see § 528.

493. Magnetic Induction. When the pole of a strong magnet is placed on the upper end of a short bar of steel which is clamped in a vertical position (Fig. 256), the steel becomes magnetic

and will support a considerable mass of iron filings at its lower end. If the magnet is now separated from the steel bar some of the filings will drop off but enough will remain to show that the steel retains considerable of the magnetism that was induced in it by the presence of the magnet. It is permanently magnetized. If a bar of soft iron of the same size is now substituted for the steel and the experiment repeated, it will be found that the iron becomes more strongly magnetic than the steel when under the influence of the magnet, but when the magnet is withdrawn it loses its magnetism almost entirely. It does not become permanently magnetized. The magnetic particles may be thought of as subject to a kind of frictional resistance in case of steel which makes their arrangement and the consequent magnetization of the steel more difficult than in case of soft iron, but also prevents the arrangement from being so easily broken up by the forces with which the molecular magnets act on each other.
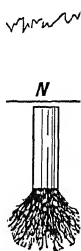
FIG. 256.
Magnetic induction

494. Magnetism Acts through Different Media. *Magnetism acts through most substances just as through air or vacuum.* Take a powerful horseshoe magnet or, better still, an electromagnet and place across its poles a thin sheet of cardboard and then bring up a mass of iron filings under the cardboard: they will cling in a great mass under the poles. If a plate of glass or lead or wood or any other non-magnetic substance be substituted they will cling in the same way. Let the plate now be fixed in position a short distance down from the poles of the magnet, as in figure

FIG. 257

257, some iron filings will fall off as it is lowered, but if not too far separated a considerable mass will still cling. Now slip in a plate *A* between the magnet and the lower plate on which the filings rest. There will be no change if the plate *A* is of wood, glass, brass or any other non-magnetic substance, but if an *iron* plate is introduced at *A* immediately some or all of the filings will fall, showing that *an iron plate will screen the region beyond, at least partially, from the action of the magnet.*
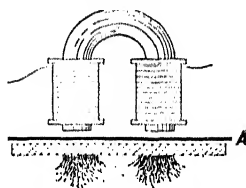
LAW OF FORCE AND MAGNETIC FIELD

495. Law of Force between Two Poles — Coulomb's Law.
The law of force between two magnets was first carefully studied
by Coulomb (1736–1806) by means of the *torsion balance*. A
magnet was suspended by a fine wire in a horizontal position
inside of a glass vessel by which it was screened from air cur-
rents. The upper end of the wire was attached to a graduated
head by which it could be
twisted through any desired
number of degrees. A second
magnet *mA* was then fixed in
a vertical position with one of
its poles near the similar pole
of the first magnet. The force
of repulsion was measured by
the torsion of the wire. By
twisting up the wire by the
head *e* the poles were brought
closer together, and the in-
creased torsion in the wire
gave the increase in repulsive
force. The method is compli-
cated by the fact that the

FIG. 258.   Torsion balance

magnetic attraction of the earth as well as the torsion of the wire
acts on the suspended magnet, and the action of both poles of
each magnet must be taken into account.   By this investigation
Coulomb was led to enunciate the law that *the force between two
magnet poles is proportional to the strength of the poles and in-
versely proportional to the square of the distance between them.*   It
may be expressed thus
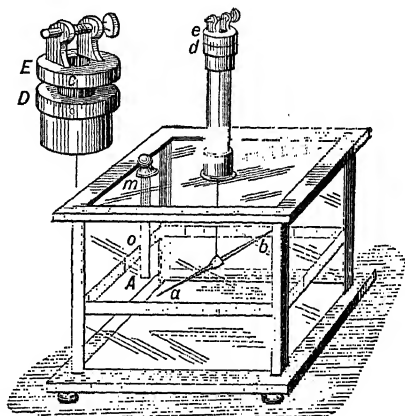
$$F = K \frac{mm'}{r^2}$$

where F represents the force, m and m' the strengths of the two
poles, and r the distance between them; K is a constant which
depends on the units in which the various quantities are meas-
ured, and, as is now known, on the medium surrounding the
magnets.   The law assumes that the poles m and m' occupy so

little space that they may be regarded as points compared with the distance r, and when so understood the most refined modern measurements only confirm its truth.

496. Field of Force. The region around a magnet is said to be a *field of force*. An interesting way of examining the field
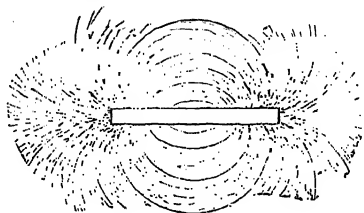
FIG. 259.   Iron filings near a magnet

of force near a magnet is as follows. Lay a sheet of glass on the magnet and dust over it fine iron filings. On gently tapping the plate the filings will gather into lines or filaments as shown in the figure. These lines indicate the direction of the magnetic force in the field. A minute compass needle placed at any point takes the direction of the line at that point. *Lines having at every point the direction in which a compass needle would stand, if placed there, are called lines of force.* Of course an in-
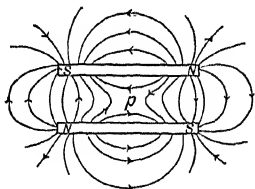
FIG. 260
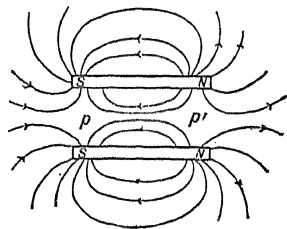
FIG. 261

finite number may be imagined drawn from one pole to the other.

In figures 260 and 261 are shown the lines of force in case of two magnets placed near each other. In figure 261 the corresponding poles are near each other and the magnets repel. It will be seen that no lines of force pass from one to the other, and two neutral points are shown at p and p'. In figure 260

the magnets are shown with the north pole of one opposite the south pole of the other. The two attract each other in this case and lines of force pass directly across from one to the other, leaving a single neutral point at p.

497. Unit Pole. For the exact study of magnetism it is necessary that certain units should be adopted as a basis for measurements.

A unit pole, or a pole having unit strength, *is one which if placed one centimeter from an equal pole in vacuum will repel it with a force of one dyne.*

It will be observed that this unit is based directly on the C. G. S. units of length and force. If these units are employed the expression for the force in dynes between two magnetic poles of strengths m and m', and r centimeters apart in vacuo, is

$$F = \frac{mm'}{r^2}.$$

For all practical purposes this expression also gives the force in air and in all other media that are not distinctly magnetic.

498. Strength of Field. When a magnet is placed in a magnetic field, due either to the earth or to some other magnet, each pole is acted on by a force which depends both on the strength of the pole and on the strength of the field in which it is placed.

The strength or intensity of a magnetic field at any point is the force in dynes on a unit magnet pole placed at that point.

Thus the earth field has a strength 0.5 at a point where a magnetic pole of unit strength is acted on with a force of 0.5 dynes.

When a pole of strength m is at a point where the strength of field is H, it is acted on by a force of Hm dynes.

In any field of force the two poles of a magnetic needle are urged in opposite directions. *The direction in which the north pole tends to move is known as the positive direction of the line of force at that point.*

499. Magnetic Moment. Suppose that in a magnetic field of strength H, a magnetic needle is placed in such a position that the line joining its poles makes an angle a with the lines of force

of the field.   Let m represent the strength, or number of units in its north pole, and $-m$ the strength of its south pole.   Then the north pole is urged with a force Hm in the positive direction of the lines of force of the field and the south pole experiences an equal force in the opposite direction.   These equal and parallel forces constitute a couple whose moment is $Hml\ sin$ a, where l is the distance in centimeters between the two poles of the magnet.   The quantities m and l belong to the magnet and their product ml is known as the magnetic moment of the magnet, and is represented by M.
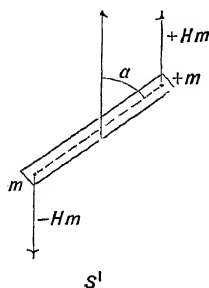
FIG. 262

Thus *the magnetic moment of a magnet may be defined as the product of the strength of one of its poles by the distance between them.*

The couple which acts on the magnet may then be expressed by the formula,

$$HM\ sin\ a.$$

500. Period of Oscillation of a Magnet in a Magnetic Field. If a magnetic needle in a field of force is disturbed from its position of rest, it will vibrate to and fro just as a pendulum oscillates in the field of the earth's attraction.   The period of one complete oscillation of a pendulum has been shown to be

$$T = 2\pi \qquad (\S\ 148)$$
$$mgh$$

where I is the moment of inertia of the pendulum and mgh is a quantity which when multiplied by the sine of the angle of inclination of the pendulum gives the moment of force which at that instant urges it toward its equilibrium position.

In case of the oscillating needle the mechanical conditions involved are the same, except that the couple causing the motion is due to magnetism instead of to gravitation.   The factor HM is the quantity which when multiplied by *sin a* gives the couple acting to turn the magnet; it, therefore, plays the same part in this case as the factor mgh in the case of the pendulum.

Hence the period of oscillation of a magnet in a field where the strength is H units is

$$= 2\pi \sqrt{\frac{I}{HM}}$$

where I is the moment of inertia of the magnet and M is its magnetic moment. It should be observed that in this case as in that of the pendulum the formula gives the period when the arc is exceedingly small. With large arcs of vibration the period is longer.

From the above formula it is clear that the stronger the field of force at a point where a magnetic needle is placed the more rapidly it will oscillate when set in vibration. This is the explanation of the rapid quivering of a compass needle when brought near the pole of a magnet.

PROBLEMS

1. A short compass needle is placed near the side of a straight bar magnet and equidistant from its poles. In what direction does it point? In what direction will it move if floating, and why?

2. Two bar magnets, exactly alike, are placed in line with each other, their north poles toward each other and south poles directed away. What is the strength of field at a point midway between the two and the direction of the lines of force near that point?

3. What is the magnetic moment of a bar magnet having poles of strength 200, and 20 cm·. apart, and what would be the moment of the couple required to hold it at right angles to the lines of force in a field of strength 5 in C. G. S. units?

4. What couple would be required to hold a magnet with pole strength 150, and with 16 cms. between poles, at an angle of 30° with the lines of force in a field of strength 2?

5. Find the ratio of the strengths of field at two places when a certain magnetic needle oscillates n times per sec. at one place and n' times per sec. at the other.

501. Strength of Field at a Point Near a Magnet. The direction and intensity of the force near a magnet may be calculated as follows. Let m and $-m$ be the strengths of the two poles of the magnet, and let r be the distance from the point P to the north pole of the magnet and let r' be its distance from the

south pole. Then if a unit pole were at P it would be subject to a force $\dfrac{m}{r^2}$ in the direction a, and to a force $\dfrac{m}{(r')^2}$ in the direction b. Laying off distances a and b proportional to the amounts of these two forces, the resultant force will be represented on
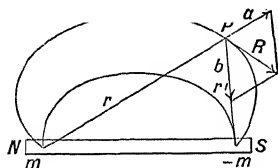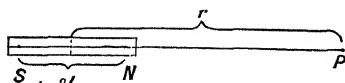
FIG. 263   FIG. 264

the same scale by the diagonal of the parallelogram on a and b. The resultant is, of course, tangent to the line of force at P.

It is sometimes desirable to calculate the force due to a magnet at a point P in line with the axis of the magnet as shown in figure 264. Let m be the strength of each of the poles, r the distance of P from the center of the magnet, and l the distance of either pole from the center of the magnet. Then by Coulomb's law the force on a unit north pole placed at P due to the north pole is $\dfrac{m}{(r - l)^2}$ and is directed toward the right. That due to the other pole is $\dfrac{m}{(r + l)^2}$ and is toward the left. The resultant force at P is then toward the right and may be written

$$F = \frac{m}{(r - l)^2} - \frac{m}{(r + l)^2},$$

or

$$F = \frac{4rlm}{r^4 - 2r^2l^2 + l^4}.$$

If l is small compared with r the terms involving l^2 and l^4 in the denominator may be neglected, as they are insignificant compared with r^4, so *approximately*

$$F = \frac{4ml}{r^3} \quad \text{or} \quad F = \frac{2M}{r^3}$$

where $M = 2ml$, the *magnetic moment* of the magnet.

PROBLEMS

1. What is the amount and direction of the magnetic force on a unit pole placed at a point in line with a bar magnet and 20 cms. away from its N pole, if the strength of the magnet's poles is 100, and its length between poles is 20 cms.?

2. The poles of a bar magnet have a strength of 200 units each and are 20 cms. apart. Find the direction and amount of the force due on unit pole to the N pole at a point 30 cms. from each pole. Find also the force due on unit pole to the S pole at the same point. Find the resultant strength of field at the point by the vector diagram.

3. Find the amount and direction of the magnetic field strength at a point 30 cms. distant from each of the poles of a bar magnet, in which the poles have strengths $+300$ and -300 and are 30 cms. apart.

4. Find the amount and direction of the force on a unit north pole placed in line with the magnet described in problem 3 and 30 cms. distant from its north pole.

5. Find the strength of the magnetic field at a point 5 cms. distant from the center of the magnet of problem 3 in a direction at right angles to its axis.

6. Calculate by the method used in § 501 the strength of field at a point on a line drawn through the center of the magnet at right angles to its axis.

TERRESTRIAL MAGNETISM

502. Declination of the Magnetic Needle. The compass needle, mounted so as to rotate in a horizontal plane, does not in general point directly north, but a few degrees east or west of north, and this deviation is called its *declination*. In observing the declination it will not do to assume that the magnetic axis of the needle is in the same direction as its axis of figure. If the magnetic axis is as represented by the dotted line in figure 265 then the apparent declination is in the first case too small. If the needle is now turned over and suspended with the opposite side upward, it will give too great an apparent declination. The mean of the two will be the true declination. This direction is called the magnetic meridian.

503. Dip or Inclination. It was observed by Hartmann (1489–1564) that if a needle were balanced before being magnetized the north end would dip downward after magnetization.

The so-called dipping needle was first made by Norman, a London instrument maker, who mounted a needle on a horizontal axis so that it could swing freely in a vertical circle. The needle was then carefully balanced so that it would stand in any position before magnetization. But after it was magnetized it was observed that the north pole pointed downward some 70° below the horizontal if the plane in which it turned was north and south by the compass.

To diminish friction the cylindrical
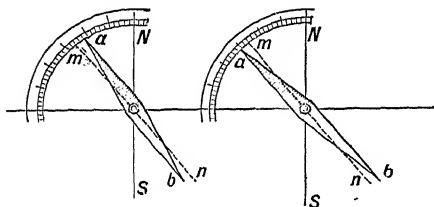
FIG. 265.   Declination of compass   FIG. 266.   Dipping needle
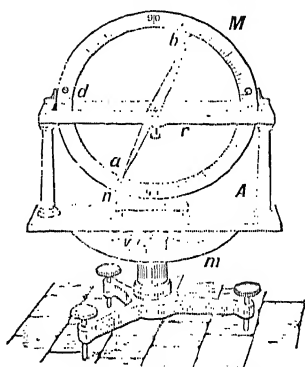
pinions on the ends of the axis of the dipping needle usually rest on horizontal plates of polished agate, on which they roll as the needle turns.

In this case also the needle must be reversed, the side toward the east being turned toward the west, to guard against error due to the axis of the needle not being in line with the direction of its magnetization. To guard against any want of balance in the needle, it should be magnetized over again with its *poles reversed*, and the dip again observed. If the needle is well constructed, the mean of these four observations will be the required dip or *inclination*.

504. Resultant Direction of Magnetic Force. The dipping needle gives the direction of the resultant magnetic force at any point. It is found that near the equator the needle is horizontal, as it is taken north its north pole points downward by an amount which increases steadily till at some point northwest of Hudson Bay it points vertically downward. That point is called the *north magnetic pole*, and *near there a horizontal compass needle would have no directive tendency and would be useless.* North

of that point the north pole of the compass needle would point south.

In vessels that change their latitude greatly the compass needle must be provided with a little sliding weight or counterpoise to correct its dipping tendency. South of the equator the south pole of the needle dips downward.

Figure 267 shows the probable form of the lines of magnetic force around the earth; of course, the direction of these lines of force is known only at the earth's surface. The magnetic condition of the interior of the earth is entirely unknown.

The declination or deviation of the resultant force from the geographic north direction also varies from point to point on the earth, at some points being east of north and at others west of north. This fact was first observed by Columbus who, as he advanced in his voyage, was alarmed to see that the compass no longer pointed as it had done when he started.
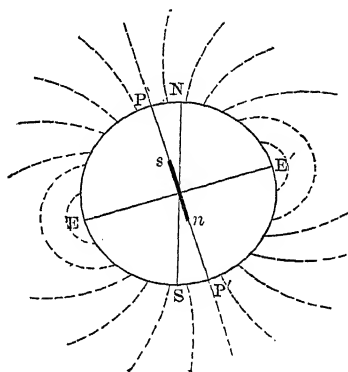
FIG. 267. Lines of force of the earth

The chart on the following page shows the declination of the magnetic needle throughout the United States in 1925. Isogonal lines connect places where the declination is the same.
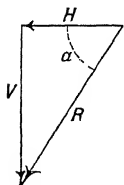
FIG. 268

505. Intensity of the Earth's Magnetism. The magnetic force of the earth at any point may be considered as the resultant of two component forces, the horizontal component H (Fig. 268), and the vertical component V. The horizontal component H is the force which is effective in directing the compass needle. The smaller this component the more feebly will the needle be affected. When a needle is balanced or suspended in the usual way so as to vibrate in a horizontal plane, its period of oscillation depends on this component only.

As shown in § 500, the intensities at different places may
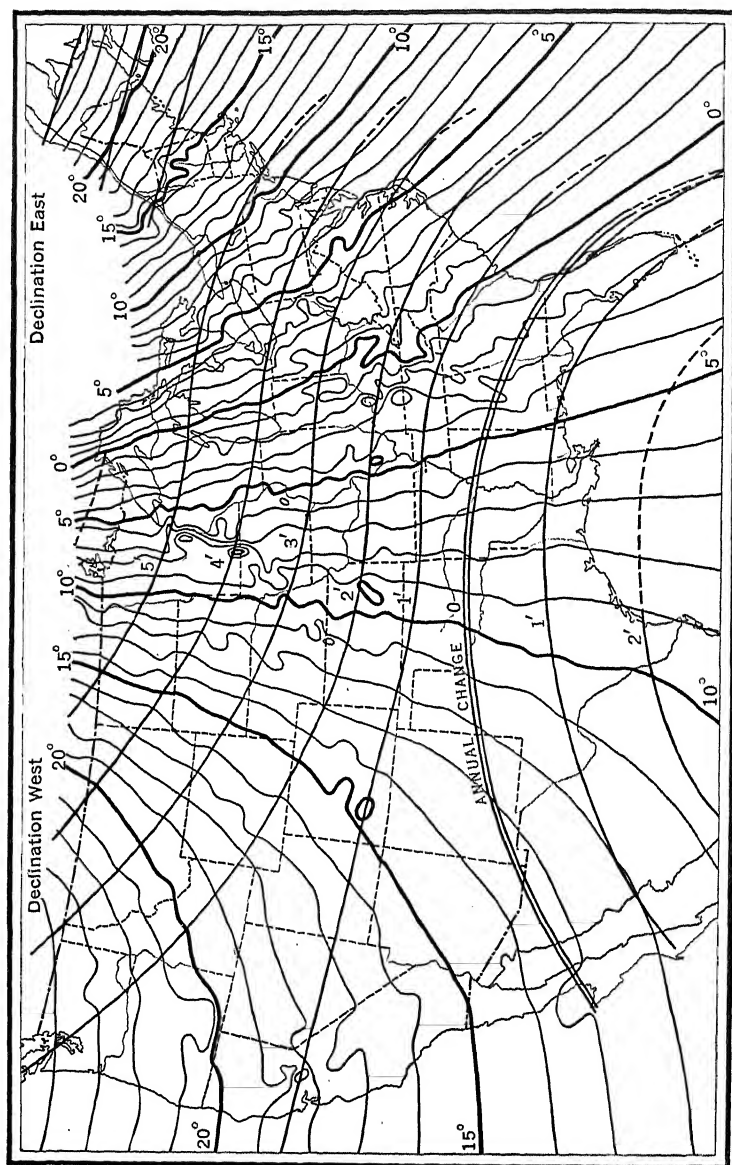
FIG. 269. Isogonal lines showing the declination of the magnetic needle throughout the United States in
(From U. S. Coast and Geodetic Survey)

be compared by causing the same magnet to vibrate first at one place and then at the other. The intensities are proportional to the squares of the number of vibrations per second. By this means the horizontal component of the intensity may be determined at any point as compared with that at some standard place.

When the horizontal intensity and the dip are both known, the resultant intensity may be found by the relation

$$R = \frac{H}{\cos a}$$

when a is the angle of dip.

The following table shows the value of the horizontal component and total force at certain places. Notice how the horizontal force becomes less in higher latitudes. The intensity is given in C. G. S. units, or the force in dynes upon a unit pole.

STRENGTH OF THE EARTH'S MAGNETIC FIELD
(DYNES PER UNIT POLE)

PLACES	DIP	HORIZONTAL INTENSITY	RESULTANT INTENSITY
Central South America.........	0°	0.30	0.30
North coast of South America..	40°	0.32	0.42
Cuba........................	50°	0.32	0.50
Georgia.....................	60°	0.26	0.52
New York...................	70°	0.18	0.53

506. Secular Change in the Magnetic Field. One of the most remarkable features of the earth's magnetism is that it is continually changing. The declination of the needle is slowly changing everywhere; that is, the magnetic poles are slowly shifting their positions. At the same time the dip is changing. The changes in declination and dip at London are shown by the curve in figure 270. The pole of a freely suspended needle would at that place apparently move through a complete cycle of change in about 470 years. This slow change is called the *secular* change.

507. Diurnal Variations. The careful study of the magnetic conditions at any place by self-recording instruments shows that there is a periodic change in the magnetic elements depending on the time of day.

The curves of figure 271 show the average variations at Kew, near London. It will be observed that the maximum changes
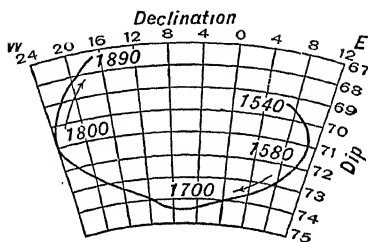
FIG. 270. Secular change in dip and declination at London
(After L. A. Bauer)

take place in the daytime and may be due to variations in temperature of the earth's surface.

508. Irregular Disturbances. The magnetic needle is also often disturbed by what are called magnetic storms; these dis-
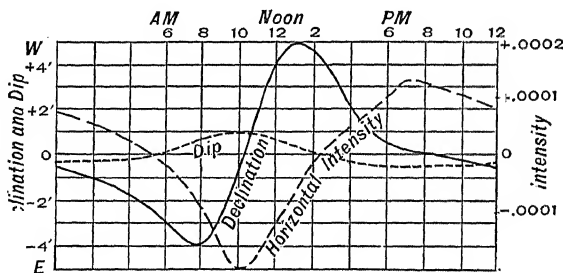
FIG. 271. Diurnal changes in dip, declination and intensity at Kew

turbances usually accompany any marked display of the *aurora borealis*, and they also seem to be more prevalent at times of sunspot maxima.

509. The Earth a Magnet. It was suggested by Dr. William Gilbert (1600), physician in the court of Queen Elizabeth and the first to take up the scientific study of magnetism, that the

earth itself was probably a great magnet, and later observations have borne out this idea. Two well-marked magnetic poles being found, one northwest of Hudson Bay in North America and the other south of Australia.

But while there is this general resemblance to a simple magnet, the direction of the magnetic force varies from place to place in a way that cannot be wholly accounted for by the supposition of simply two poles.

The magnetism of the earth seems to be due to a variety of causes, the presence in the earth of magnetic masses is a cause of local variations and may have great influence in the surface layer of the earth, but it seems probable that the temperature in the interior of the earth is too high for it to possess any very strong magnetism. Electric currents flowing in the surface of the earth and due to its varying temperature as first one side and then another is exposed to the sun, as well as currents of electricity in the upper air, probably play an important part in determining its magnetic state. But the complete explanation has not yet been given, and any theory to be satisfactory must account for the remarkable secular changes in its magnetism which go on slowly and progressively year after year.

510. Gauss' Method of Measuring the Horizontal Intensity. The horizontal component of the earth's magnetic force may be measured by the following method due to Gauss. A small steel bar magnet is suspended horizontally by a fine fiber in a closed box by which it is protected from air currents. It is then set oscillating through a small arc and the period of oscillation carefully determined. This period depends on M the magnetic moment of the magnet and on H the horizontal component of the earth's magnetic force. By § 500

$$HM = \frac{4\pi^2 I}{T^2}$$

where I is the moment of inertia of the magnet, a quantity that is determined by its mass, size, and shape, and T is the period of a complete oscillation. The product HM is thus found.

To determine the relation of H to M a second experiment is necessary.

Suppose P is the point where the magnetic force H is to be determined and where the period of oscillation of the magnet NS was observed in the first experiment. Place at P a very short magnetic needle, while the magnet NS is placed exactly east or west of P and with its axis on the east and west

line, as shown in figure 272. If r is the distance from the center of NS to P, then the force at P due to the magnet is, as shown in § 501,

$$F = \frac{2M}{r^3}.$$

Then at P the force H due to the earth and the force F due to the magnet are at right angles to each other, as shown by the arrows in the figure. The

Fig. 272

needle at P will take the direction of the resultant force R and will therefore be deflected through the angle a, but

$$tan\ a = \frac{F}{H} \quad \text{or} \quad \frac{2M}{r^3 H}$$

whence

$$\frac{H}{M} = \frac{2}{r^3\ tan\ a}$$

This expression shows that to determine the ratio of H to M it is only necessary to measure the distance r and the angle of deflection a.

Having by the first experiment determined the product HM and by the second the ratio $\frac{H}{M}$, it only remains to multiply the two together to find H^2 and so determine H.

Unit Tubes of Force

511. Number of Lines of Force. Up to this point lines of force have been regarded as simply expressing the direction of the force in the magnetic field. We must now follow Faraday in a very remarkable development of the idea.

In a stream of water flowing steadily lines may be imagined drawn which at every point are in the direction of flow, and which may be called stream lines. An infinite number of such lines may be drawn. The whole stream may then be conceived to be divided up into *tubes of flow* by means of surfaces which everywhere coincide with stream lines. These tubes of flow may be taken of such a size that each will transmit the same

quantity of water per second, say one cubic foot. Then, where the stream is most rapid, the cross sections of the tubes of flow will be smallest and they will widen out as the velocity diminishes. The whole number of such tubes in the stream will be equal to the number of cubic feet of water transmitted per second. These tubes of flow may be called *unit tubes*, and the number of them crossing perpendicularly a surface 1 sq. ft. in area is equal to the number of cubic feet of water crossing that area per second. Thus *the number of unit tubes passing perpendicularly through a unit surface at any point in the stream is equal to the velocity at that point.*

Now in the same way the magnetic field may be conceived as divided up into unit tubes by means of surfaces parallel to the lines of force. And it may be proved that where such a unit tube is smaller the field is more intense, and where it widens out the strength of field is less, just as the velocity varies in case of the stream of water. So that it is possible to take these tubes of such a size that *the number passing perpendicularly through a square centimeter of surface at any point may be equal to the strength of the magnetic field at that point.*
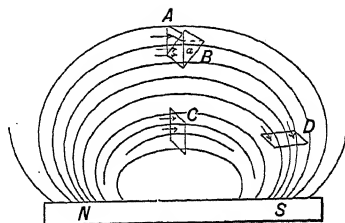
FIG. 273

We may imagine that *each unit tube is represented by a line of force drawn through its center or axis, and when the phrase number of lines of force is used it refers to such lines.*

Using the term in this way, it is clear that in the case shown in figure 273 more lines of force pass through a card in position C than in position A, as the force is greater at C than at A, and consequently there are more lines of force to the square centimeter. Clearly, also, fewer lines of force pass through the card in position B than in A and most of all in position D. If the card were placed parallel to the lines of force none at all would pass through it.

The number of lines of force through A is found by taking the average strength of the field at the surface A and multiplying this by the area of A in square centimeters, since the number of lines per square centimeter is equal to the strength of the field at that

point.   If the surface is oblique to the lines of force as at B, the number of lines of force passing through it will be found by multiplying the number in the perpendicular position A by the *cosine* of the angle a, or, what comes to the same thing, multiply the average strength of the field at the surface B by the projection of that surface on a plane at right angles to the lines of force.

PROBLEMS

1. How many lines of force pass through a square meter of floor area where the total strength of the earth's magnetic field is 0.6 and the lines of force are inclined 60° from the horizontal?

2. How many lines of force in this case would pass through an area of 1 square meter on an east and west wall, and how many in case the wall ran north and south?

3. How many lines of force pass through an area of 4 sq. cms. placed as at A in figure 273, with its center 12 cms. from each pole of a magnet 20 cms. long between poles, strength of poles being 288?

4. How many lines of force pass through a circle 1 cm. in diameter placed 8 cms. from the north pole of a bar magnet 16 cms. long, which points directly at it and has poles of strength 200?   The plane of the circle is perpendicular to the axis of the magnet.

512. Lines of Force Inside a Magnet.   The lines of force of a magnet are not to be supposed as only outside of it.   If we
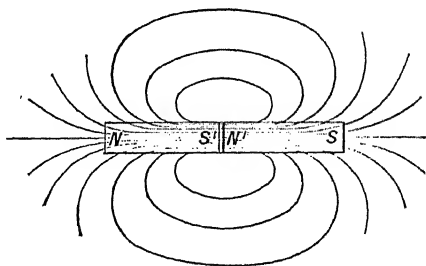
Fig. 274

imagine a minute magnetic needle placed in a crack extending across the magnet it will be acted on most powerfully by the poles $N'S'$ on each side of the crack, but it will also be affected by the attraction of the end poles N and S of the magnet.   In consequence of the superior influence of the poles $N'S'$, it will set its

north pole toward the left or S'. If we now imagine the cleft shifted along toward the north end of the magnet, the force inside the cleft will become less because it will be nearer to N, which pole tends to make the needle point in the opposite direction. But still the needle will point from right to left. When the cleft comes infinitely near to the end N, the magnetism of N and of S', which form two opposite and equally magnetized layers, will neutralize each other so that the effect is the same as though the needle were just outside the magnet at N. We see in this way that *the force in the cleft is absolutely continuous with that outside of the magnet:* there is no abrupt change in passing through the surface. *The force in such a cleft is called the magnetic induction and the lines of force outside of a magnet form continuous closed curves with the lines of induction inside of the magnet.* The lines of force outside are also called lines of induction, as there is no distinction between the two except inside of a magnetic medium. What is called the positive direction of these lines is from the north to the south pole *outside* of the magnet. Of course as many lines of force as emerge from the north pole enter at the south pole, and all the lines of force or induction in the magnet pass through its middle section. Looked at in this way, *the poles are seen to be simply those regions where the lines leave or enter the magnet, and the most intensely magnetized portion of the magnet is the center where the lines of induction are closest together.* If a little block could be cut from the center of the magnet without disturbing its magnetism, it would be found a more powerful magnet than a similar block cut from any other part where the lines of induction are not so close.

WARNING: In using the words *entering* and *emerging* with reference to lines of force nothing like *flow* or *motion* must be supposed; when what we arbitrarily call the *positive direction* of the line of force is toward a surface, it is spoken of as entering it; and when that direction is away from a surface the line of force may be said to leave the surface or emerge from it.

513. Influence of the Shape of a Magnet on Its Power and Retentiveness. A short thick bar of steel is more difficult to magnetize strongly than a long thin one and loses its magnetism more easily. A thick magnet may be thought of as made up of a bundle of thin ones of the same length. But it is clear that in such a bundle each little magnet would tend to set up lines of

force down through its neighbor in such direction as to oppose or weaken the other's magnetism.

Thus *there is a demagnetizing tendency which is greatest in a short thick magnet.* Horseshoe magnets are long and have their poles close together and consequently there is very little demagnetizing tendency. There is, however, a tendency for the lines of force in this case to pass across on the inside of the poles instead of out at the ends. A soft-iron block placed across the poles, and called an armature or keeper, provides an easy path for the lines of force from one end around to the other and thus tends to keep the poles near the ends.
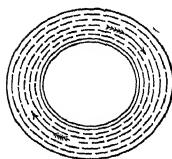
FIG. 275

514. Ring Magnet. A uniform ring of iron or steel may be magnetized by means of an electric current so that the lines of force are circles entirely within the substance of the ring. In such a case the magnet has no poles as there are no places where the lines of force enter or leave the ring. Such a magnet has no external field of force and would not act on a magnetic needle placed near it, and yet it is magnetized, as will be evident if it is broken, for in that case each half will show two poles.

MAGNETIC INDUCTION

515. Induction Studied by Iron Filings. If the lines of force of a horseshoe magnet are examined by means of iron filings on a plate of glass, as described in § 496, and if a bar of soft iron is then placed a short distance in front of the poles of the magnet and the field again examined in the same way, a notable change will be observed. The lines of force are bent toward the two ends of the soft-iron bar as though they could be established in the iron more easily than in the surrounding medium. And the softer the iron and the more easily it is magnetized, the greater the number of lines of force that will pass through it rather than the more resisting medium around it. Thus the presence of the iron makes the field of force weaker beyond it, and the nearer the iron bar is to the poles of the magnet the more lines of force will be drawn into it and the fewer there will be in other parts of the field.

516. Permeability. *The ratio of the number of lines of force established in any medium to those established in a vacuum by the*

*same magnetizing force has been called by Lord Kelvin the permeability of the medium.** From a physical point of view, permeability of a medium may be thought of as the ease with which the lines of force are established in that medium as compared with a vacuum. Thus iron has a permeability several hundred times greater than air. Most other substances have a permeability which is sensibly the same as air or vacuum, and, therefore, the magnetic field is practically the same in wood, glass, or water as in air.

A hydraulic analogy may aid in forming a clear conception of this subject. Imagine a stream of water continually flowing out of the north of the horseshoe magnet (Fig. 276) and entering its south pole. Suppose the medium surrounding the magnet was of a uniform porous nature that opposed considerable resistance to the flow from *N* to *S*. The lines along which the flow would take place would be like the lines of force in the field before the soft iron was introduced. Imagine a cavity to be made in the porous medium having just the size and position of the soft-iron bar. Lines of flow now would tend toward this cavity through which the liquid would flow freely and a correspondingly smaller flow would take place in other regions.
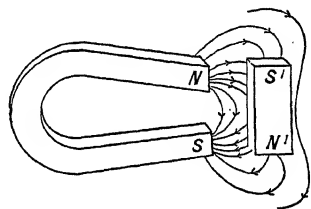
FIG. 276

The lines of flow in this case correspond to the lines of force when the soft-iron bar with its great permeability is in the field.

517. Magnets Formed by Induction. When a soft-iron bar is placed in front of a magnet as shown in figure 276, at the end nearest the north pole of the magnet the lines of force are directed toward the end of the bar as toward the south pole of a magnet and at the other end they are directed away from the bar as from a north pole. The bar of iron thus becomes a *magnet by induction*. If it were of steel it would retain some of this magnetism when taken out of the field.

Suppose a *long* bar of soft iron to be placed in a magnetic field

* The meaning of this can be better understood by imagining a narrow axial cleft perpendicular to that shown in the magnet of figure 274. The lines of induction must all cross the transverse cleft, but they will avoid the axial cleft as much as possible and follow the iron path instead. *Permeability* is measured by the ratio of the field intensity in the transverse cleft to that in the axial cleft.

parallel to the direction of the lines of force. The result will be
as shown in figure 277, lines of force will be drawn into the bar
in consequence of its great permeability entering it at one end and
leaving it at the other, so that one end becomes a south pole and
one a north pole. On each side of the bar the field is weakened.

When, however, the bar is placed across the field of force as in
figure 278 it will have only a very slight effect on the field of force
since the lines of force can pass through only a small thickness of
iron. So also a thin flat sheet of iron placed perpendicular to the
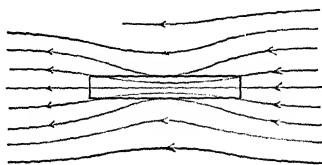
FIG. 277.   Bar of soft iron parallel   FIG. 278.   Bar of soft iron
with lines of force of field   across the lines of force
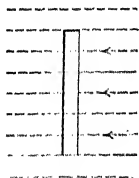
lines of force of the field would have practically no effect on the
field.

**518. Effect of Heat and Jarring in Case of Magnetizing by
Induction.** The magnetism induced in an iron or steel bar
placed in a magnetic field parallel to the lines of force may be
increased by striking the bar with a hammer or jarring it while
under the influence of the field, also by heating the bar red-hot
and allowing it to cool in the magnetic field. These disturbances
seem to facilitate the arrangement of the molecules under the in-
fluence of the magnetic force and help to overcome the resistance
to magnetization which especially characterizes hard steel.

519. Magnetic Induction in the Earth's Field. If a bar of
soft iron having no permanent magnetism is placed in the earth's
field parallel to the lines of force, that is, in the direction of the
dipping needle, its lower end in north latitudes will become a
north pole and its upper end a south pole, as may be shown by a
magnetic needle. If jarred by the blow of a hammer while in this
position it will be found permanently magnetized. If, however,
it is placed at right angles to the lines of force of the earth it is
scarcely magnetized at all (§ 517).

In consequence of induction iron ships are magnetized by the earth differently when pointing in different directions.

In such vessels the standard compass is usually *compensated* by having soft-iron bars so placed near it that the magnetism induced in them will in every position just balance that induced in the ship, while permanent steel magnets may be used to compensate the permanent magnetism of the ship.

520. Hysteresis. When the magnetic field in which a mass of iron is placed is varied in strength, the changes in the magnetism of the iron lag behind the changes in the field. This is known as hysteresis and is discussed in connection with the magnetization of iron by electric currents, § 692.

PERMEABILITY, DIAMAGNETISM, AND INFLUENCE OF MEDIUM

521. Magnetic Substances Attracted. *When a fragment of iron is placed in a magnetic field it experiences a force in that direction in which the strength of the field increases most rapidly.*
If at A (Fig. 279), it is drawn directly toward the magnet in the direction of the lines of force. If at B, it is drawn toward the magnet at right angles to the lines of force. If at C, it will be drawn in a direction oblique to the line

FIG. 279

of force somewhat as shown. If it is in a uniform field, as in the earth's magnetic field, or is at a point in a magnetic field where the force is a maximum or a minimum, it will be in equilibrium and have no tendency to move in any direction. Such a point of equilibrium would be found midway between two equal poles either like or unlike.
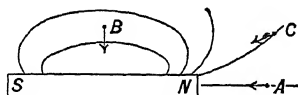
If the fragment is long in shape it will turn and *point* in the direction of the line of force, but it will not always tend to *move* along that line.

Any substance whose permeability is greater than vacuum will act in this way in a vacuum and such are known as *paramagnetic* or simply *magnetic* substances.

522. Diamagnetic Substances. Faraday (1845) experimented on the behavior of a great variety of substances in the intense field between the poles of a powerful electromagnet. A little

oblong of pure copper when suspended by a fine fiber in this field was found to set itself at right angles to the lines of force, as shown in figure 280. So also fragments of wood, paper, aluminum, bismuth, glass, and many other substances. These substances Faraday called *diamagnetic*. Substances like nickel, cobalt, and manganese which behave like iron, setting themselves in the direction of the lines of force, he called *paramagnetic* or *magnetic*.
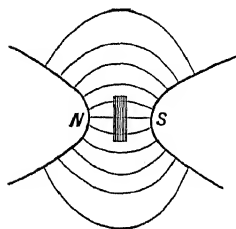
FIG. 280. Bismuth in magnetic field

Diamagnetic substances when placed in a magnetic field are driven from a stronger field toward a weaker, *the force acting on a fragment of such a substance being in the direction in which the strength of the field diminishes most rapidly.* This may be well shown in the following way. A ball of bismuth, which is the most strongly diamagnetic substance known, is suspended between the poles of a powerful electromagnet, being hung from one end of a light arm of wood which is itself supported in horizontal position by a delicate bifilar suspension, so that the slightest force will cause the arm to swing around carrying the ball out of the magnetic field. If while the ball hangs between the two poles the current is applied to the electromagnet, the bismuth ball will at once be driven aside out of the intense field.

The setting of the diamagnetic bars *across* the lines of force described at the beginning of this section finds its explanation in the preceding experiment; for the field of force between the magnet poles is most intense next the poles as is shown by the crowding together of the lines of force, and so the ends of the bar are in a much less intense field when the bar stands across the lines of force than if it were to be directed along them; it therefore assumes the former position.

523. Influence of the Medium. By the following interesting experiment Faraday showed that *the medium surrounding a body in a magnetic field plays an important part in determining the magnetic force upon it.*

When a thin-walled glass capsule, long in shape, is filled with a weak solution of ferric chloride and suspended between the poles of a magnet, it sets itself along the lines of force showing that

the ferric chloride is *magnetic*. This happens whether the capsule is hung in air or water. *If, however, it is surrounded by a solution of ferric chloride stronger than that within the capsule it will act as if diamagnetic,* placing its length across the lines of force.

524. Permeability of Magnetic and Diamagnetic Substances. When the permeability of a substance is greater than that of the surrounding medium, the lines of force are drawn in toward the substance, as already discussed in § 517 and as shown in figure 281 which represents the disturbing effect of a ball of substance whose permeability is greater than that of the medium around it.

If, however, the permeability of the ball is less than that of the medium, the lines of force will be spread, as shown in figure 282. A magnetic needle placed near the ball will point aside instead of toward it.

In the first case if the ball is in a field that is not uniform, as near the pole of a magnet, it will be attracted or drawn toward the stronger field. If, however, the ball has a permeability less than the surrounding medium, it will be driven away from the pole toward a weaker field.

Magnetic or paramagnetic substances may then be defined as those whose permeability is greater than that of vacuum, while those whose permeability is less than vacuum are diamagnetic.
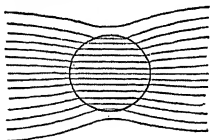
FIG. 281. Permeability of ball greater than that of medium
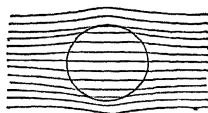
FIG. 282. Permeability of ball less than that of medium

525. Magnetism of Gases. Faraday also studied the magnetic qualities of different gases. Oxygen gas was found to be attracted toward the poles, while hydrogen was repelled. Oxygen was thus shown to be more permeable than air. Later experiments have shown liquid oxygen to be decidedly magnetic.

526. Magnetic Alloy. In 1903 Heusler made the very interesting discovery that an alloy of 25 parts manganese, 14 aluminum, and 61 copper, had decided magnetic properties, although none of the substances of which it is made is magnetic except

in the very slightest degree. It seems to indicate that magnetism depends upon molecular rather than atomic structure. The permeability of this alloy has been found to be nearly 33.*

527. Effects of Heat on the Magnetism of Metals and Magnets. The permeability of iron and nickel diminishes as the temperature rises. At 737° C. iron ceases to be magnetic. A small piece of iron heated to a bright red heat is not attracted even by a powerful magnet, but as it cools to 700° it again becomes strongly magnetic. A steel magnet when heated to bright red heat loses all trace of magnetism, and if cooled while away from magnetic influence will be found completely demagnetized.

Even when a magnet is slightly heated, say to 100° C., it is not as strong as at lower temperatures.

528. Ferromagnetism. Nearly all of the elements are diamagnetic or paramagnetic to a slight degree. Iron, cobalt and nickel, however, stand entirely by themselves because they are many times more magnetic than any of the other elements. These three elements, which are closely associated in the periodic table of elements, are called the *ferromagnetic* elements.

Magnetism is undoubtedly a property of the atoms themselves. It will be seen later that an electric current flowing in a coil makes the coil a magnet (§ 688), with a north pole at one end and a south pole at the other, determined by the direction of flow of the electric current. The atoms with their revolving electrons (§ 801) should therefore be minute magnets, but in ordinary substances these magnets are pointing in random directions as shown in figure 254, so that no magnetism is exhibited. When a magnetic field is applied, the motions of the electrons are slightly changed, so that in some cases the substance becomes slightly paramagnetic and in others it becomes diamagnetic depending upon the electron arrangement in the atom itself.

In ferromagnetism the case is different. Here, the application of the magnetic field causes an effect as though the entire atoms were rotated and all made to point in one direction as shown in figure 255 so that all of these minute magnets act to-

---

* Recently an alloy of 78 parts nickel and 22 parts iron called *permalloy* has been developed by the Western Electric Co., which has a high permeability in weak fields, so that even the earth's magnetic field may produce in it a high degree of magnetism.

gether and build up a very intense magnetization. Permalloy (§ 526) is the most perfect example of a ferromagnetic substance known. Hysteresis (§ 692) is a characteristic of ferromagnetism. Permalloy exhibits far the smallest hysteresis of any known ferromagnetic substance. All ferromagnetic substances are paramagnetic.

529. Force with which a Magnet Attracts its Armature. The force with which a magnet attracts its armature evidently depends on the fact that the permeability of the armature is greater than that of the surrounding medium. *If there were no difference between them there would be no change in the lines of force on withdrawing the armature and consequently no attractive force.*

When the armature is of such a size that most of the lines of force from one pole to the other pass through it, the force of attraction is given very nearly by the formula

$$P = \frac{AB^2}{8\pi}$$

where A is the combined area of the two poles and B is the *induction* or the number of lines of force that pass from a pole into the armature across a square centimeter of surface. If these quantities are taken in C. G. S. units, the attractive force P will be found in dynes.

PROBLEMS

1. Find the force 5 cms. away from a pole of strength m, the other pole being so far away in comparison that it may be disregarded. How many lines of force go through the sphere of 5-cm. radius surrounding the pole m?

2. If a magnet having poles of strength 300, and 30 cms. apart is mounted on a pivot in a uniform magnetic field of strength 0.2, how much force, applied 10 cms. from the pivot, will be required to hold it at right angles to the lines of force of the field?

3. What is the magnetic moment of the magnet in problem 2? What torque is required to hold it at an angle of 45° to the lines of force of the earth field of strength 0.2?

4. Where the total intensity of the earth's magnetic field is 0.6 and the dip 70°, how many lines of force pass through a circular hoop 50 cms. in diameter lying horizontally on the floor? How many if the plane of the hoop is vertical facing north and south?

5. If a compass needle oscillates 2 times per sec. when 15 cms. distant from the pole of a long magnet, how fast will it vibrate when 8 cms. from the pole, neglecting the influence of the other pole?

ELECTROSTATICS

530. Electrification. If a hard-rubber rod is rubbed with fur or flannel it will attract light fragments of pith, paper, or gold leaf. A light ball of pith suspended by a thread, as shown in figure 283, is strongly attracted. A rod of sealing wax, or sulphur, or indeed of dry wood, will show the same power. In cold dry weather if a piece of paper is laid on a table and rubbed with flannel, it will be found to cling to the table and a slight crackling may perhaps be heard as it is pulled away; and if in this condition it is held near the wall it will be drawn toward it and cling to it. The shavings which come from a carpenter's plane in winter when the air is very dry will often behave in the same way. In all of these cases the substances are said to be *electrified*, a term which comes from *electron*, the Greek word for *amber*, a substance which was known to the ancients to possess this power.
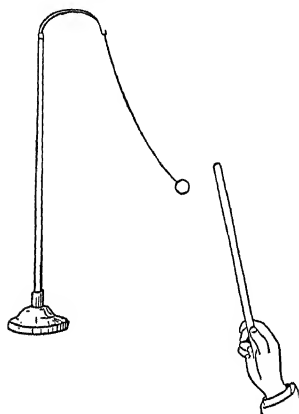
FIG. 283

About the year 1600, Dr. Gilbert, who was also a pioneer in the study of magnetism, found that a very large number of substances could be electrified by rubbing, though with metals he could get no results. He accordingly classified substances as *electrics* and *non-electrics*, according as they could, or could not, be electrified by rubbing.

531. Conductors and Non-conductors. In 1729 Stephen Gray discovered that the electrification of a glass rod would leak off from it and could be communicated to a ball through a damp cord. His experiments showed that electrification could be communicated through certain bodies which he called *conductors* while it could not be communicated through others which were named *non-conductors*.

Metals were found to be the best conductors, wood and damp cord were fairly good, while glass, sulphur, and resin were non-conductors.

Gray then showed that the substances which Gilbert had classed as non-electrics were conductors, and if they were *insulated* or mounted on non-conducting supports they could be electrified as other substances can. The old distinction of electrics and non-electrics was therefore abandoned, and substances were classified as *conductors* and *non-conductors or insulators.*

The insulating power of bodies may be compared by the time required for a given amount of electrification to leak through similar rods of the different substances.

In the following table bodies are classified according to their *resistances* or insulating powers.

Insulators	*Poor Conductors*	*Good Conductors*
Amber	Dry wood	Metals
Sulphur	Paper	Gas carbon and graphite
Fused quartz	Alcohol	Aqueous solutions of salts
Glass	Turpentine	and acids
Hard rubber	Distilled water	
Air and gases		

532. Electricity. Take two metal pails, each mounted on an insulating support (Fig. 284), electrify one of them and then connect the two by a conductor, such as a cord or a wire. Both will show electrification when tested by the suspended pith ball, though the electrification of the one first charged will be less than before the two were connected.

(The connecting conductor must be supported on glass

Fig. 284

rods or from loops of silk thread or otherwise insulated when placed on the pails.)

If the pails are connected by a metallic wire the redistribution of the electrification is instantaneous; if by a rod of wood or by a cord a perceptible time is required.

The communication of electrification from one body to another,

one always losing as the other gains, suggests a transfer of *something* of which electrification is the external evidence. *This something is called electricity, and when electrification is communicated from one body to another, there is said to be a flow of electricity.*

533. Two Kinds of Electrification. Rub a rod of hard rubber or sealing wax with fur, and when strongly electrified present

Fig. 285

it to a suspended pith ball. The ball will be attracted at first, but if allowed to touch the electrified rod it may cling for a moment and then spring away, strongly repelled.
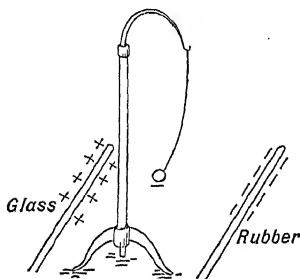
If a rod of glass, electrified by rubbing with silk, is now brought near the pith ball, it will fly to the glass, but after contact it will be repelled as it had been from the rubber rod.

While repelled by the glass it will be attracted by the electrified rubber, and *vice versa.* It is clear that the electrical states of the glass and rubber are different.

This discovery was made by Du Fay, a French investigator, in 1733. He found that all electrified substances behave either like glass or rubber, and the two kinds of electrification were accordingly called *vitreous* and *resinous.* Franklin named the electrical state of the glass *positive* and that of the rubber *negative*, and these names have been universally adopted.

534. Similarly Electrified Bodies Repel. Two strips of hard rubber electrified by fur and suspended near together repel each other. Two strips of paper if drawn through a fold of flannel in dry cold weather, or even when drawn between the fingers, will repel each other and stand apart. On the other hand, a strip of rubber negatively electrified by fur is attracted by a rod of positively electrified glass.

Similarly electrified bodies repel, while oppositely electrified bodies attract each other.

535. Electron Theory. The phenomena of electrification which have been just described are simply and satisfactorily explained on the basis of the *electron theory.* According to this theory the atoms of all matter (§ 797) consist of central positively

charged nuclei surrounded by negatively charged particles called *electrons*. In the neutral state, when no electrification is present, the negative charges of the electrons which surround the atoms are sufficient to neutralize the positive charges of the atomic nuclei. Although the electrons individually repel each other, the positive charges of the nuclei are sufficient to hold them in their proper places in the atom. Certain of the outer electrons of the atoms are not very strongly held, however, and are easily displaced or removed. These external electrons are often called the *valence electrons* because they determine the chemical valence of an atom. When a body is positively charged it has lost some of its negative electrons. When it is negatively charged it has an excess of electrons clinging to it. Thus when the rod of hard rubber is rubbed with fur as described in § 530 a certain number of valence electrons are transferred from the fur and to the rod making the latter negative and leaving the fur positively charged. On the other hand in the case of the glass rod, electrons are removed from it and cling to the fur making it negative and leaving the glass positive.

536. Electric Series. *In every case of electrification by friction the substances rubbed together become oppositely electrified*, because electrons are taken from one and added to the other.

When glass is rubbed with silk the glass becomes positively charged and the silk negatively, but when hard rubber is rubbed with silk the rubber becomes negative and the silk positive. The silk thus becomes negative in one case and positive in the other. And, in general, *any substance may become either positive or negative, depending on what it is rubbed with*. It is possible to arrange substances in a series such as the following in which any substance is more positive than those below it in the list, but is negative to those that precede it.

Glass (surface rubbed clean and polished).
Fur.
Flannel.
Glass (passed through a Bunsen flame).
Silk.
Wood.
Sealing wax.
Hard rubber.
Sulphur.

537. Gold-leaf Electroscope. An instrument, such as the suspended pith ball, used to detect electrification is called *an electroscope*. A much more sensitive electroscope is that shown in figure 286.

Two strips of thin gold leaf about $\frac{1}{2}$ in. wide and 3 in. long are attached to the end of an insulated brass rod so that they hang side by side in a glass jar which screens them from air currents. The brass rod passes through the insulating stopper of the jar, and terminates above in a plate or knob.
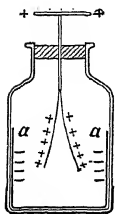
FIG. 286

Two strips of tinfoil *aa* on the inside of the jar are in metallic connection with a metal base or tinfoil coating over the outside of the bottom.

If a charge is given to the upper plate of the electroscope it at once distributes itself by conduction over the rod and gold leaves, causing the latter to repel each other and diverge as shown in the figure. If the charge should be too great the leaves will diverge enough to touch the side strips *a* through which the whole charge will escape.

538. Conduction. When electricity is conducted through a body the outer loosely held electrons of the atoms are passed on from atom to atom in a continuous stream. At the same time the conductor may remain electrically neutral because as soon as an atom passes an electron to its neighbor on one side it receives a new one from its neighbor on the other side. Electric currents are thus made up of negative electricity. Undoubtedly, had the pioneer investigators appreciated this fact, an electric current would have been considered as flowing in the direction of electron motion and not in the opposite direction as it is according to the present convention. For purposes of analysis it makes no difference which way the current is thought of as flowing. It should be kept in mind, however, that the positive direction of flow of electricity, as ordinarily understood, is opposite to the actual electron motion.

A conductor is a material which easily passes its electrons from atom to atom in this way.

A non-conductor or insulator is a material in which the electrons are not readily passed from atom to atom in a continuous

stream, although, through applied friction, valence electrons may be removed easily or extra ones may become attached, as described in § 535.

Law of Force and Distribution of Charge

539. Coulomb's Law. The French physicist, Coulomb (1784), investigated the law of force between two electrified bodies using a torsion wire balance, illustrated in figure 287. By means of a very fine wire a light hori-

zontal bar of shellac, glass, or other insulating material is suspended inside of a glass jar by which it is screened from air currents. On the end of the suspended bar is a light pith ball n which is covered with gold foil. A metal ball m mounted on the end of an insulating glass rod can be introduced into the glass jar through a hole in the cover.

To use the instrument remove the ball m and by means of the graduated head from which the wire is suspended turn the wire until the ball n hangs exactly in the place of m. Now give a charge to m and introduce it into the jar. At first n is attracted and touches m, the charge then divides between the two since both balls are conductors, and immediately n is repelled to such a distance that the twist in the wire balances the force of repulsion. The distance between the balls is observed and also the

Fig. 287. Coulomb's balance

number of degrees through which the wire is twisted as read on scale C in the case.

Now increase the twist in the wire by means of the graduated head, thus forcing n toward m. It will be found that when the two are at one-half the first distance the force of repulsion as measured by the twist in the wire is four times as great.

To study the effect of changing the quantity of charge, a second insulated brass ball is taken of the same size as m and mounted on a glass rod in the same way. The ball m with its charge is now withdrawn from the jar and touched to the other similar ball which has no charge. Immediately the charge divides equally between the two (since they are alike) and m now has only half the charge which it had at first. If it is carefully introduced

without permitting it to touch n, the charge on the latter will not be changed, and if the force is observed when the balls are the same distance apart as in the first experiment it will be found that the force is only one-half as great.

From many such experiments Coulomb concluded that *the force between two given charged bodies, provided they are small compared with the distance between them, is inversely proportional to the square of the distance and directly proportional to the amounts of their charges.*

This law may be expressed algebraically thus:

$$F = \frac{qq'}{Kr^2}$$

where F represents the force, K is a constant, and r is the distance between the centers of the two bodies whose charges are represented by q and q'.

The constant K depends on the units that may be used, and also, as was shown by Faraday, on the medium between the two charged bodies.

540. Unit Charge. *Unit charge or unit quantity of electricity, in the electrostatic system of units, is defined as that quantity which when placed one centimeter from an equal charge in vacuum repels it with a force of one dyne.**

The force in dynes between two electric charges in vacuum may therefore be expressed by the formula

$$F = \frac{qq'}{r^2}$$

where the quantities q and q' are measured in electrostatic units and where r is the distance between the charges, measured in centimeters.

When the charges are in any other medium the force is usually less and the formula is .

$$F = \frac{qq'}{Kr^2}$$

* The number of electrons in unit charge equals 2.1×10^9 or over two thousand millions.

where K is a constant usually greater than one, known as the *specific inductive capacity* or *dielectric constant* of the medium.

The force between two charges in air is appreciably the same as in vacuum, for the specific inductive capacity of air is greater than that of vacuum by only about one part in 2000.

541. Distribution. The distribution of an electric charge may be examined by means of a little metal disc mounted on an insulating handle and known as a *proof plane*. If the disc is placed flat against the surface of the charged and insulated pail shown in figure 288 and then removed, it will carry away a charge which may be tested by the gold-leaf electroscope. When examined in this way it is found that the greatest charge
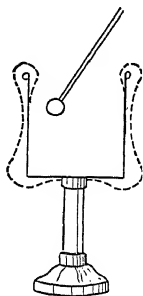
FIG. 288. Distribution on pail
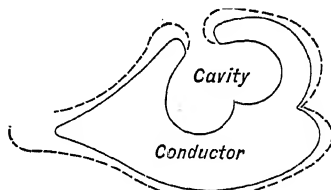
FIG. 289. Density of distribution indicated roughly by dotted line

is obtained from the outer surface of the pail near its upper edge and on the outer corner at the bottom, less is found on the middle of the side and none at all in the interior except near the upper edge.

When there is a metal cover on the pail, absolutely no charge can be found on any part of the interior.

Other irregular bodies may be examined in the same way and it will be found that the greatest density of charge is at corners and knobs projecting outward. For example, in a conductor shaped as in figure 289 the greatest density will be found on the projecting point on the left, no charge will be obtained in the cavity even though there is a sharp point there, and very little will be found toward the bottom of the dimple on the right end.

542. Charge Entirely on the Surface of a Conductor. The following experiment was carried out independently by Cavendish and Biot.

A metal ball having two closely fitting hemispherical metal cups which were provided with insulating handles, was insulated and then charged strongly with electricity. *When the cups were simultaneously removed they were found to have the entire charge,*
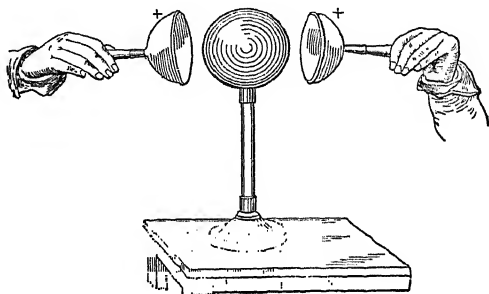
FIG. 290

the ball being left without any trace of electrification, showing that the whole charge was on the surface.

543. Discharge from Points. The density of charge on sharp projecting points is much greater (§ 541) than it is on other parts of a conductor. When this density exceeds certain limits the air in the neighborhood is made slightly conducting by a collision process described in § 769. Some of the molecules of the gas become positively charged and others negatively. Those having a charge opposite to that on the point are attracted to it and gradually discharge it while those having a charge of the same kind as the point are repelled and carry away as much charge as the point has lost. The repelled charged molecules, by collision with the neutral molecules of the air, cause a stream of air to flow from the point, which is called the "electric wind." If the point is connected to an electric machine this wind may be strong enough to blow out a candle placed near the point or turn a little wheel having light vanes. If the point from which the discharge takes place be movable, it will be driven backward by the repulsion of the charge in the air and a continuous rotation

may be produced by the arrangement of points shown in figure 291. Conductors which are designed to hold electrical charges should therefore have all projecting parts or corners carefully rounded, otherwise they will be rapidly discharged.

544. Frictional Electric Machine. The early forms of electrical machines were frictional; the one illustrated in figure 292 is a good type of this class. A circular glass plate is mounted firmly on an axle so that it can be turned between leather covered rubbers, which are pressed against the glass by springs. The rubbers remove electrons from the glass by friction, leaving the glass positive. The positive glass plate then passes between a pair of branches extending from a metallic conductor which is on an insulating support of glass or of hard rubber. On the inside of each of these branches is a row of sharp metal points projecting toward the glass plate like the teeth of a comb. Electrons are drawn off of these points on to the positive glass plate by the proc-

FIG. 291. Electric wind

ess described in the last section, making it neutral again but leaving the metallic conductor positively charged because of its loss of electrons. The action is continuous as the plate revolves,
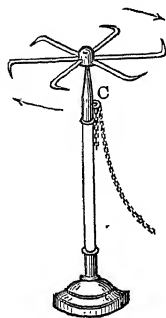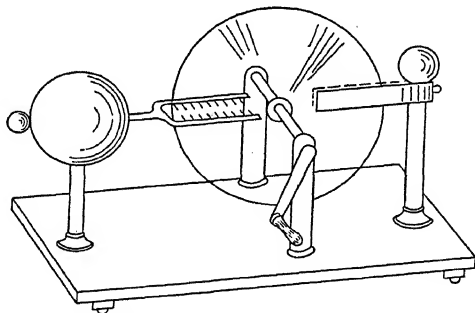
FIG. 292. Electrical machine

the metallic conductor becoming charged positively more and more while the rubbers acquire an increasing negative charge. It is advantageous to connect the rubbers to the ground by a chain or wire to permit its excess of electrons to escape.

PROBLEMS

1. If the charges on two small conducting pith balls are $+8$ and -8, will they attract or repel and with what force when 4 cms. apart? What if they are allowed to touch?

2. Two small conducting balls of the same size and 6 cms. apart have charges $+36$ and -4, respectively. What is the force between them? Also what will the force become if they are touched together and then placed as before?

3. Two pith balls 3 cms. apart and equally charged repel each other with a force of 16 dynes; find the charge on each.

4. Two pith balls hung from the same point on threads of equal length and weighing $\frac{1}{10}$ gm. each, are equally charged and repel so that they diverge until the threads are at right angles to each other. What is the force of repulsion in grams and in dynes? If they are 8 cms. apart, what is the charge on each?

5. Two pith balls, each weighing $\frac{1}{10}$ gm. and suspended from the same point by threads 30 cms. long, are equally charged and repelling each other, hang 8 cms. apart. What is the charge on each ball?

INDUCTION

545. Induction. When a conductor having no charge is insulated and then brought near a positively charged body, such as A, figure 293, it is found to be negatively electrified on the side next to A and positively electrified on the farther side. This can readily be tested by means of the proof plane and electroscope.

In this case no electrons have passed from A to B. The conductor B is insulated and no flow of electrons either into it or out of it is possible. When B is taken away from the charged body A it is found to have no charge, just as before it was brought up. Also no charge is lost or gained by the body A in the operation. The charges produced in the conductor B by the proximity of A are said to be *induced*.

546. The Induced Charges are Equal. If, instead of the one conductor B of the previous experiment, we take *two* insulated conductors B and C, having no charge, and bring them near the charged body A while in contact with each other, as shown in figure 294, a negative charge will be induced in the nearer one and a positive charge in the farther one, that is, negative

electrons drawn toward the conductor A accumulate on B making it negative and leaving C with a positive charge. If they are now separated and then removed, the displaced electrons are trapped on one leaving the other one positive. On bringing them near an electroscope they are found oppositely electrified;
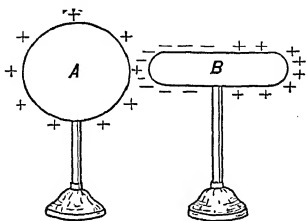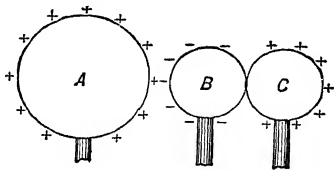
FIG. 293   FIG. 294

and if while away from the vicinity of A they are touched together, the charges totally disappear, showing that the negative electrons of the one were just sufficient to neutralize the positive charge of the other. *The charges are therefore equal and opposite.*

The same result is reached whatever the shape or size of the conductors B and C. Thus B may be large and C small, or *vice versa*, and still the charges on each when removed from the influence of A will be found to be equal and opposite.

For if we consider the single conductor B used in the first experiment (Fig. 293) it is clear that if it is imagined cut in two at any point, near the positively charged end for example, the positive charge on the smaller part will be equal to the excess of negative over positive on the greater part.

Thus *the positive and negative charges produced by induction are always equal.*

547. Induction when the Conductor is Already Charged. When the conductor has a charge to begin with the inductive action takes place in the same way, but the original charge is combined with the induced charge. Thus if B (Fig. 295) is given a positive charge and brought toward the positively charged body A, it will be found that the positive charge is denser on the side away from A.
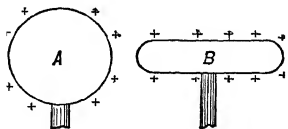
FIG. 295

At a certain distance there will be no charge at all on the end toward A. If brought still nearer a negative charge will be found on that end, while the positive charge on the rest of the conductor will be correspondingly increased.

If A and B are both conductors and similarly charged, say positively, they will react on each other, and the greatest density will be found on the outer side of each.

If, however, they are oppositely charged the greatest densities are on the adjacent sides.

548. Effect of Connecting with the Earth. If, while in the presence of the positively charged body A, the conductor B is

connected with the earth by a wire, or by touching it with the finger, the induced charge on B is supplemented by electrons attracted from the earth by the positive charge on A, so that the free positive charge on B is completely neutralized and B now has an excess of electrons.
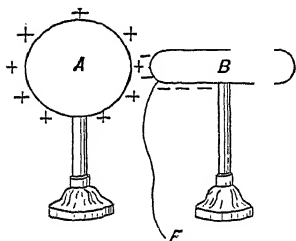
It makes no difference what part of B may be connected to the earth, whether the nearer end, as shown, or the farther end, the result is exactly the same.

FIG. 296

The induced charge that remains is sometimes called a *bound charge* because it does not flow off or disappear when B is touched, but is held by the presence of the charged body A.

549. Charging Electroscope by Induction. When a rod of hard rubber, negatively electrified by rubbing with fur, is brought toward a gold-leaf electroscope, the leaves will be observed to diverge strongly while the rod may be several inches from the instrument.

Electrons from the top of the electroscope move downwards into the leaves, due to the repulsion of the rod, thus giving them a negative charge precisely as in the case discussed in § 545.

If the top of the instrument is touched by the finger while the electrified rod is still held near, electrons will escape, being driven off by the repulsion of the rod until the leaves hang straight down as in b (Fig. 297). If the finger is now removed and then the rod, the remaining electrons will redistribute themselves

leaving an excess of positive charge on both the top and on the leaves of the electroscope and the leaves will diverge as shown in *c*.

When the electroscope is positively charged the approach of a positively charged rod increases the divergence of the leaves, while a negatively charged body will draw them together unless it is brought too near, in which case they will again diverge with a negative induced charge.

It will be noticed that there are shown in figure 297 induced charges on the metal side strips inside the electroscope. These
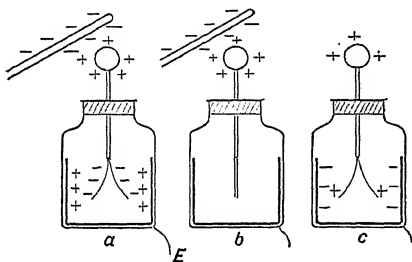
FIG. 297
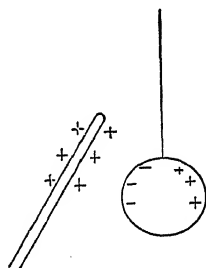
FIG. 298

charges are opposite to the charge on the leaves and increase their divergence.

550. Attraction of Pith Balls Explained by Induction. When an electrified rod is brought near a pith ball (Fig. 298) an inductive action takes place as shown, and the attraction between the positively electrified rod and the negatively electrified side of the ball is greater than the repulsion of the positively electrified side, since the negative side of the ball is nearer to the rod.

If the ball had an initial positive charge it would be repelled, though even in this case if the rod is much more strongly electrified than the ball and is brought very near to it there may be attraction.

551. Induction Takes Place through Non-conductors. The interposition of a sheet of glass or hard rubber or a cake of beeswax or any other insulator between an electrified body and an electroscope does not interfere with the inductive action. Indeed induction takes place more readily through these substances

than through air, though but slight evidence of change would be observed in such a rough experiment.

552. Induction through Conductors. If a charged rod (Fig. 299) is held over an electroscope and a small *insulated*
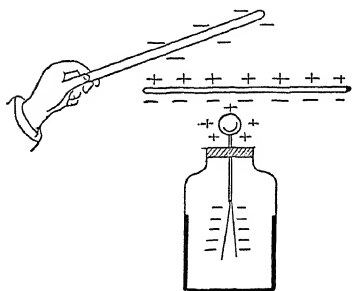
FIG. 299

sheet of metal is interposed between the two, the gold leaves will diverge as though the plate were not there, for if the rod is negative a positive charge will be induced on the upper side of the plate and an equal negative charge on the lower side which in turn will act on the electroscope.

If, however, the plate is connected with the earth, electrons will escape from it leaving its lower surface with no charge, and the electroscope will be screened almost entirely from the effect of the rod.

553. Conductor Surrounding an Electroscope. When an electroscope is entirely surrounded by a conducting sheet it is absolutely protected from all *outside* inductive action. It has already been shown (§ 541) that there is no electrification on the interior of a closed conductor, so also there is no induction from the outside. If a delicate electroscope is enclosed in a cage of wire gauze which is underneath as well as around and above it, the cage may be strongly electrified by a machine and there will be no disturbance of the electroscope, except such as may be due to electrified air passing into the interior. Faraday constructed a small room about 6 ft. each way and covered with tinfoil and found that within it he was unable to detect any disturbance of his most delicate electroscope though an assistant was electrifying the outside by a machine so that long sparks escaped from it.

554. Electrophorus. A simple form of induction apparatus devised by Volta is known as the electrophorus. It consists of a cake or plate of non-conducting material, such as resin, sulphur, or hard rubber, supported on a metal base and having a metal cover provided with an insulating handle of glass or hard rubber. The upper surface of the resinous plate is negatively

electrified by rubbing it with fur and the cover is then placed upon it. Electrons are driven from the lower to the upper side of the cover by repulsion, producing a negative charge on its upper side and leaving the lower side with an equal positive charge. On touching it with the finger or connecting it by a metal chain or wire with the base plate the excess of electrons escapes from the upper side, leaving it with no charge but leaving the positive charge on the lower side as before. If the cover is now lifted from the resin, on presenting the knuckle, electrons return to it again in a bright spark thus neutralizing the positive charge. The cover may then be placed again on the resin, touched, and withdrawn and a second spark obtained and so on indefinitely. In this

FIG. 300. No electric force within enclosing conductor

way a great number of charges may be obtained without renewing the electrification of the resinous sole plate.
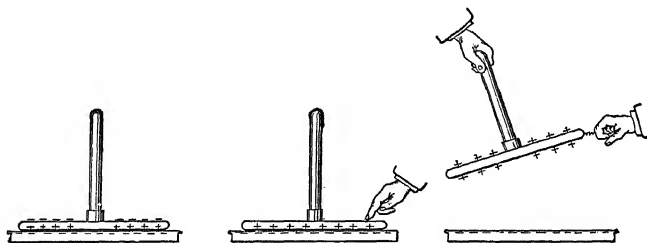
FIG. 301. Electrophorus

The resin is a good insulator and the cover touches it at so few points that there is very little direct loss by conduction.

555. Source of the Energy of the Charges. Each of the charges obtained in this way has energy, as shown by the noise

and light given out by the spark. This energy does *not* come from the energy spent in electrifying the plate of rubber or resin, for a spark is obtained every time the insulated cover is touched and withdrawn without any appreciable loss of electrification by the resin.

The energy must be supplied in the operation of withdrawing the plate. This will be made evident by the following experiment.
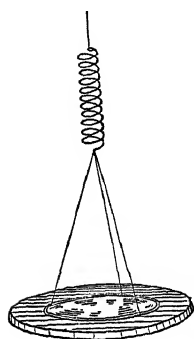
Suspend the cover by silk cords from a spring, and after having discharged it let it be lowered upon the resin and then withdrawn *without being touched*. The spring is scarcely stretched more when the plate is withdrawn than when it was lowered. But if when the cover is on the resin *it is touched*, the negative charge escapes and the attraction between the positive charge in the cover and the negatively charged resin causes the spring to be greatly stretched when the plate is raised, showing that more work has to be done to raise the plate after it has been touched than before.

FIG. 302

It is this work done by the person lifting the plate that is the source of the energy of the charge that is obtained.

If the cover is raised only an inch from the resin the spark will be much less energetic than if it had been raised 10 in., for less work has been done.

556. Faraday's Ice-pail Experiment. A very important case of induction is that of a charged body surrounded by a conductor. This was first investigated by Faraday as follows:

Taking an insulated metal pail having a metal cover and connected with an electroscope, it was observed that when a charged metal ball was brought up toward the pail the divergence of the leaves of the electroscope increased *until the ball was entirely within the pail, after which no change was observed whatever the position of the ball might be,* whether it was close to the bottom or to one side or in the middle.

The ball was now permitted to touch the inside of the pail, but not the slightest change in the gold leaves was observed. When the ball was withdrawn it was found completely discharged while the leaves remained diverging.

The same observations may be made using a deep open pail, as in figure 303, provided the ball is not too near the open top.

When the positively charged ball is introduced into the pail there is induced a negative charge on the inside of the pail leaving a positive charge on the outside, as may be shown by a proof plane.

When the ball touches the interior of the pail the charges on the ball and on the interior of the pail disappear, for the ball and pail then become one conductor and there is no charge on a cavity in a conductor.

If these charges were not exactly equal there would be some excess of either positive or negative charge which would cause a flow of electrons from the outside to the inside or vice versa and produce a charge in the electroscope.

The experiment then leads to the following conclusion: *When a charged body is surrounded by a conductor a charge is induced on the inside of the conductor equal and opposite to that on the body.*

FIG. 303

The walls, ceiling, and floors of ordinary rooms are fairly good conductors so that when we have a positively charged body in a room we may be sure that an exactly equal negative charge is distributed over the walls and neighboring objects.

557. Positive and Negative Electrifications Always Equal. Hold a rod of sealing wax in an insulated pail connected with an electroscope and rub it with a pad of flannel which is insulated on another rod of sealing wax.

They may be rubbed quite vigorously but no sign of electrification is shown by the electroscope, but if either one by itself is drawn out of the pail there is decided divergence of the gold leaves. It follows that the electrifications developed on the sealing wax and flannel, respectively, are equal and opposite. This indeed is exactly what is to be expected on the basis of the electron theory as discussed in § 535.
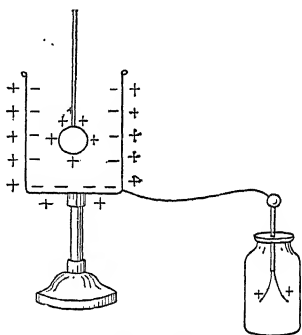
In every case of electrification, whether by induction or friction, equal positive and negative charges are produced.

558. Induction Machines. Various forms of electrical machines have been devised in which charges developed by induction from small initial charges are continually added to the original charges until powerful effects are obtained. The first powerful and successful machine of this kind was made by Holtz about 1864.

A modification of this machine, due to the labors of Voss and Toepler and known as the Voss-Holtz or Toepler-Holtz machine, is shown in figure 304 with a diagram illustrating its action.

A circular plate of glass carrying on its front surface six small discs of tinfoil, marked $a_1, a_2, \ldots a_6$, is rotated rapidly
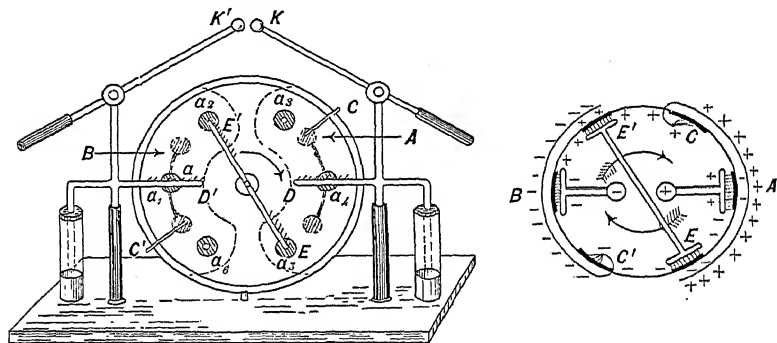
FIG. 304.　Toepler-Holtz machine and diagram

in front of a fixed plate of glass, on the back of which are attached two conductors of paper A and B, called armatures (outlined by the dotted lines). In front of the rotating plate are mounted on insulating supports the two conducting combs DD' with sharp points close to the plate and directed toward it; these conductors are connected to the knobs K and K'. A conducting bar, called the equalizing bar, EE', crosses diagonally in front of the rotating plate and is also provided with combs directed toward the plate.

At C and C' there are metallic arms which are connected with the armatures on the back of the fixed plate and carry little tinsel brushes that touch the tinfoil discs on the revolving plate as they pass.

Suppose, now, that A is slightly more positive than B, owing to the remains of a previous charge or to the influence of some neighboring charged body or to the brushes at C and C' rubbing a little differently on the discs as the plate is turned. And suppose that the discs a_2 and a_5 are under the combs of the equalizing bar EE' and are connected with it by the tinsel brushes carried by that bar. The bar with the two discs thus forms one continuous conductor with the two inductors A and B opposite its ends. Negative electrons will therefore be drawn over toward A, leaving an equal positive charge on the end toward B. If the plate is now turned, a_5 carries its negative charge past the position a_6 until it is between the brush C' and the armature B with which that brush is connected, and is therefore situated almost as if inside of a conductor; it accordingly gives up almost its entire charge through the brush C' to the armature B which thus becomes more negative. At the same time the disc a_2 has moved past the position a_3 and has received electrons again from the armature A through the brush C, leaving A more positive.

The armatures with their increased charges act more powerfully on the next pair of discs that pass under the bar EE', and so these discs carry still larger charges to the armatures and thus the effect rapidly increases till the armatures are so highly charged that they lose by leakage as rapidly as they gain.

When the armature A is positively charged it acts inductively on the comb D through the two layers of glass, drawing electrons off of the points of the comb onto the surface of the revolving glass plate and thus leaving the knob K strongly positively charged. On the other hand the inductive action of the negatively charged armature B results in a leakage of electrons from the glass plate onto the sharp points of the comb D', giving the knob K' a strong negative charge and leaving the glass plate with a positive charge, and if the gap between K and K' is not too great, spark discharges will take place between them. Two small Leyden jars (§ 581) are connected with the conductors D and D' and act as reservoirs in which the charge accumulates between discharges.

559. Wimshurst Machine. In the Wimshurst machine two circular plates of glass are revolved in opposite directions, one in front of the other.

On the outer surface of each are a number of radial strips or sectors of tinfoil, on each of which is a little metal knob or button. As the plates rotate, the buttons strike the tinsel brushes of a pair of equalizing bars, one of which is fixed in front of each plate, one inclined to the right and one to the left, so that the two are nearly at right angles to each other.

In the diagram the inner and outer circles represent the two plates while the heavy lines indicate the positions of the conducting sectors.

Suppose that the plates turn in the directions of the arrows and that the sector a is slightly positive and b negative.

Then c and d which are connected by the equalizing bar will become oppositely charged by induction, c negatively and d
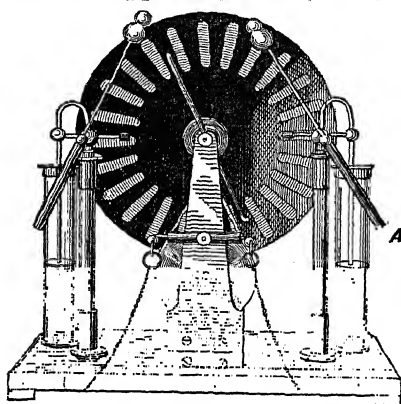
FIG. 305. Wimshurst machine   FIG. 306. Diagram
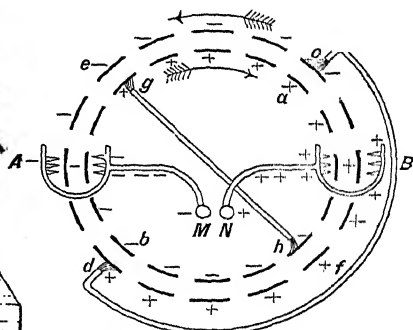
positively. The rotation of the plates carries c with its negative charge to e, and d with its positive charge to f, where they are opposite the ends of the second equalizing bar and by induction electrons are drawn over to h making it negative and leaving g positive.

At A and B are combs between which the plates turn, the rotation of the plates carries the positively charged sectors toward B and the negatively charged ones toward A, making the knobs N and M positive and negative respectively.

560. Measurement of the Charge of the Electron. The amount of charge of the electron was first measured by Sir

J. J. Thomson in 1898. Since that time this elementary charge has been measured in several different ways, but the most beautiful method of all was developed by R. A. Millikan in 1910, by which minute electric charges on microscopic drops of oil spray from an atomizer were accurately determined.

A drop was isolated and observed through a low power microscope as it slowly settled down through air in the space between two horizontal metal plates which were connected together so that there was no electric force in the region between them. When the drops had nearly reached the lower plate the two plates were electrified, one positive and the other negative, in such a way that the electric force on the charged drop carried it upward. As it neared the top the plates were once more connected and discharged permitting the drop to settle again — and so on indefinitely.

It was possible in this way to observe a single drop for hours at a time, and to measure accurately the velocity with which it settled downward and also the velocity with which it was urged upward. From the former of these two measurements the size of the drop could be determined, and then from its upward velocity in the electric field its charge could be calculated. *In several thousand such experiments the charges upon the drops, whether positive or negative, were always found to be exact multiples of a small charge e, which had the value* 4.77×10^{-10} *in electrostatic units.*

All electric charges whether large or small are a whole number of times this elementary charge so that it is impossible to increase or diminish an electric charge by a fractional part of e.

561. Electron Theory of Matter. These experiments of Millikan together with researches relating to electric discharge in gases, and in the field of radioactivity, show that all electric charges whether positive or negative are exact multiples of the unit charge e having the value 4.77×10^{-10} electrostatic units, as given in the previous section, which is so small that there are more than 2000 million of them in the electrostatic unit as defined in § 540.

The *positive* elementary unit charge is never found separate from atoms of matter, but the *negative* unit, as was shown by Sir J. J. Thomson, is carried by the small particles or corpuscles

that make up the cathode rays in a vacuum tube (§ 777) and have only $\frac{1}{1840}$ the mass of the hydrogen atom. These negative particles are the electrons which take part in the electrostatic phenomena described in the previous sections. They exist in all kinds of matter, can pass through conductors (§ 538) and may be transferred from one body to another.

The electron theory supposes that every atom of matter in the neutral state is made up of a certain number of elementary positive units and an equal number of electrons.

When an atom loses an electron it becomes positive, when it gains an extra one it becomes negative.

562. Summary. The following is a summary of the main preceding facts relating to electric charges.

1. There are two kinds of charges. Bodies with like charges repel and with unlike charges attract each other.

2. The force between two small charged bodies is directly proportional to their charges and inversely proportional to the square of the distance between them.

3. The force between charged bodies also depends on the medium between them.

4. All bodies in the neutral state contain positive and negative electricity in equal amounts.

5. The appearance of charge is a result of the displacement of negative electricity, or electrons.

6. Whenever a positive charge disappears an equal negative charge also disappears.

7. The *total charge* in a body, or the sum of the positive and negative charges which it may exhibit, does not change so long as the body is truly insulated.

8. The distribution of charge in a conductor is influenced by neighboring charged bodies. (Induction.)

9. Charges are always multiples of an elementary unit e taken a whole number of times.

It is seen that several of these statements follow directly from the electron theory of charge.

563. Other Theories of Electricity. The early investigators thought of electrified bodies as containing something which they called an *imponderable fluid*, because it could flow from one body to another and yet did not seem to possess weight or inertia.

Symmer conceived two such fluids, positive and negative electricities, which neutralized each other when mingled.

Franklin, however, advocated the view that there was but a single electricity and that for every body there was a normal amount when it showed no electrification; if there was an excess it showed one kind of electrification, while if there was a deficiency of the electric fluid the body was electrified in the opposite way.

This theory is seen to be similar to the electron theory in that it explains how opposite charges neutralize each other and how it is impossible to produce a negative charge anywhere without at the same time causing an equal positive charge to appear somewhere else.

POTENTIAL AND ELECTROMETERS

564. Potential or Electrical Pressure. Suppose an electroscope connected with a charged pail by a wire. It makes no difference in the indication of the electroscope whether it is connected to the inside of the pail where no charge can be obtained by the proof plane or to an edge where the density is greatest. There is perfect equilibrium between the electroscope and the charged pail and no tendency of the charge to flow from one to the other. When two conductors are in

FIG. 307

this relation they are said *to be at the same potential* or *to have the same electrical pressure.*
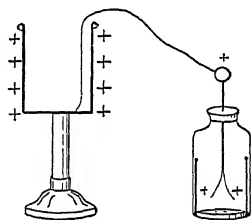
The potential of a conductor is that electrical condition which determines the flow of electricity.

Potential determines the flow of electricity just as pressure determines the flow of fluids and temperature the flow of heat.

When any conductor is connected with the earth, flow takes place until the conductor comes to the potential of the earth.

When two charged conductors are connected by a wire, the one that loses positive charge is said to have been at a higher potential than the other.

565. Effect of Increased Capacity. If an insulated conductor having no electric charge is brought up and touched to the charged pail shown in figure 307, part of the charge will flow into the conductor and the whole system comes to a new state of equilibrium in which the potential is less than before. This

is shown by the fact that the gold leaves of the electroscope do not diverge so strongly.

In this case *there has been a decrease in potential though there has been no change in the total amount of the charge.* The enlarged system of conductors is said to have a greater electrical *capacity* than the original pail and electroscope.

Change in potential due to a change in capacity of the charged conductor is well shown by Faraday's apparatus, figure 308, in which a roller suspended by insulating silk cords carries a conducting ribbon of tinfoil. When rolled up and charged the pith balls diverge widely, but as the ribbon is unrolled, thus increasing the surface and capacity of the conductor, the pith balls approach each other. That this is not due to any loss of charge is shown by the fact that when the ribbon is again rolled up the pith balls diverge as at first.
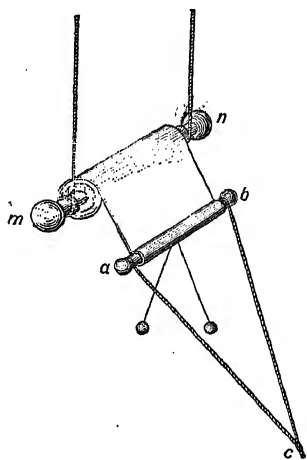
566. Potential of a Conductor. In non-conductors the potential may be different at different points, but in a conductor or in connected conductors, when the electrons are at rest, all parts are at the same potential, since otherwise flow would take place from one part to another.

Fig. 308

Even when a conductor is hollow the potential in its interior, due to electrons at rest, is everywhere the same as at its surface, provided it does not contain any insulated charged bodies; for Faraday showed (§ 553) that there was no electric force inside of a hollow conductor in such a case and consequently it must be a region of uniform potential.

Thus the surface of a conductor is an *equipotential surface*, since all points of it are at the same potential, and the interior of a conductor, provided it does not contain insulated charged bodies, is an *equipotential region*.

567. Zero Potential. Bodies are usually discharged by connecting them with the earth, and its potential is accordingly taken as the *zero*. Bodies which take on electrons, when con-

nected to the earth, thereby losing a positive charge, are said to have a positive potential while bodies which lose electrons, when connected to the earth, thereby becoming more positive, have a negative potential.

The walls and floors of wood and plaster which enclose ordinary rooms are conductors, though they conduct rather slowly, the interiors of rooms are therefore to be regarded as cavities in conductors which are at the earth potential. *When there are no insulated charged bodies inside of such a room it is an equipotential region, all at zero potential, even though there may be electrified clouds floating overhead.*

568. Potential Without Charge. It was shown by Faraday that when an insulated conductor is touched to the inside of a hollow conductor which completely surrounds it, it receives absolutely no charge. It follows that in such a case there is no flow of electrons from one to the other and therefore *both must have been at the same potential before they touched.* We see, then, that merely putting a conductor without charge inside of a hollow conductor brings it to the potential of that conductor; that is, *a conductor which has no charge takes the potential of the region where it is placed.*

569. Potential Affected by Neighboring Charges. The case just discussed is a special instance of the general principle that *the potential of a conductor depends not only on its own charge, but on that of all neighboring objects.*

This is well shown in case of the electrophorus (§ 554), for when the metal cover of that instrument is removed and discharged by touching it, it comes to the earth potential. But when it is placed on the negatively charged base its potential is lowered, as is shown by the fact that if it is now connected with the earth *electrons flow from the cover to the earth.* It thus comes to the earth potential and has a positive charge.

If it is now removed from the negatively charged base plate its potential is raised, for on touching it *it receives electrons from the earth.*

It thus appears that *the potential of a conductor depends not only upon its own charge, but also upon all other charges near enough to affect it.*

If a conductor were removed from all other charged bodies

its potential would depend only on its own charge and would be proportional to that charge. But if a positively charged body is brought near the conductor its potential is raised though its charge is not changed.

A positive charge not only raises the potential of the body to which it is given, but it raises the potential of the whole neighboring region and of any bodies that may be near. So also a negative charge lowers the potential of all points in its vicinity.

570. What Determines the Potential of a Conductor. From what precedes it will be seen that the potential of a conductor depends upon the following three conditions:

1. *Its capacity — determined by its size and shape.*
2. *Its charge.*
3. *The charges on surrounding bodies.*

571. Measure of Difference of Potential. When a small electric charge is moved along an equipotential surface or from one part of an equipotential region to another, *no work is done, for there can be no electric force acting on the charge since there is no tendency to flow.*

If two conductors are at the same potential no work is done in transferring a small charge from one to the other. But if they are not at the same potential work must be done to carry a small positive charge from the one at the lower potential to the one at the higher, just as in case of two vessels, each containing a fluid under pressure, work must be done by a pump to force any fluid from the vessel in which the pressure is less into the one in which it is greater.

It may be proved that if a little charge is transferred from one conductor to another the work done will be the same by whatever path the transfer may be made, and accordingly the work done may be used as a measure of the difference of potential of the two conductors.

Thus *the difference of potential of two conductors is measured by the number of ergs of work required to transfer unit charge from one to the other.* Difference of potential determined in this way is in electrostatic units and one *electrostatic unit of potential is very nearly equal to 300 volts*, the *volt* being the unit of potential in what is called the *practical system* of units.

572. Instruments to Measure Potential. *Electroscopes and electrometers are potential measuring instruments.* For instance, the deviation of the gold leaves of a gold-leaf electroscope depends on the difference of potential between the leaves and the side conducting strips that are connected with the earth. The leaves have the same potential as any conductor that may be connected with them. But this instrument, in the ordinary
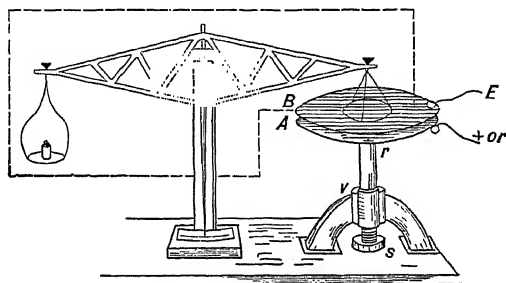
Fig. 309. Attracted-disc electrometer

form, is not well adapted for exact measurements, though a modified form in which the deflection of a single narrow strip of gold leaf is measured by a low power microscope, is valuable for some investigations.

573. Attracted-disc Electrometer. The instrument shown in figure 309 is known as the *attracted-disc electrometer*, or *the Kelvin absolute electrometer*, and may be used to measure large differences of potential.

Two circular flat plates of metal, A and B, are mounted parallel to each other. By means of the screw s the plate A may be raised or lowered and the distance between the two plates may be determined by the scale and vernier v. The upper plate is made in two parts, a central disc and a surrounding ring. The disc is suspended from one arm of a balance and hangs so that its lower surface is exactly flush with that of the ring, which is separately supported, though the two are in conducting communication.

The plate B is connected with the earth, while A, which is insulated by hard rubber or glass at r, is connected with the charged conductor whose potential is to be determined. There

is a charge on A and an opposite induced charge on B and consequently an attraction between them. By means of the balance the force with which the disc is attracted is exactly measured.

The difference of potential V between the plates A and B may then be found in electrostatic units by the formula

$$V = d \sqrt{\frac{8\pi F}{S}}$$

where d is the distance between the plates in centimeters, F is the force of attraction in dynes, and S is the area of the disc in square centimeters.

The instrument is called an *absolute* electrometer, because its determinations depend directly on measurements of length and force and it may be used to standardize other instruments.

The *guard ring*, as it is called, which surrounds the attracted disc was introduced by Lord Kelvin to cause a uniform distribution over the central disc, without which the difference of potential could not be calculated by the above simple formula. For in case of two parallel plates the distribution is denser toward the edges, but is extremely uniform near the center if the plates are not too far apart.

The balance must be enclosed in a metal case, as shown by the dotted lines, to screen it from all outside disturbing electrical attractions.

574. Quadrant Electrometer. The quadrant electrometer, also designed by Lord Kelvin, is shown in figure 310. A small round brass box is cut into four quadrants which are slightly separated from each other, and mounted on insulating supports as shown in the figure. The needle consists of a thin flat plate of aluminum, broad at the two ends as shown in the plan, and mounted on a light vertical wire of aluminum which passes through its center and carries on its upper end a small mirror by which the motions of the needle are observed.

The flat needle is suspended by a fine quartz fiber or by two parallel fine fibers of silk constituting a bifilar suspension, so that it hangs horizontally in the middle of the box formed by the four quadrants and in the position shown.

The diagonally opposite quadrants AA are connected by wire

DIELECTRICS

395

conductors to the pole P', while the quadrants BB are connected to the pole P. The needle is given a positive charge so that if the A quadrants are connected with a positively charged body while the B quadrants are joined to earth, it will turn toward the B quadrants; while if the A quadrants are negative it will turn toward them. The deflection of the needle is read by the motion of a narrow beam of light, reflected from the attached mirror upon a graduated scale, and is nearly proportional to the difference of potential between the A and B quadrants.

The sensitiveness of this instrument may be many times greater than that of the gold-leaf electroscope. To secure the greatest sensitiveness a very light paper needle is used, hung by an exceedingly fine quartz fiber.

Another method of using the instrument is to connect the B quadrants to the body to be tested, while the needle and A quadrants are connected together and to the

FIG. 310. Quadrant electrometer

earth. In this case the needle will turn toward the B quadrants whether the charged body is positive or negative, and the deflection is nearly proportional to *the square* of the difference of potential measured.

ELECTRON THEORY AND DIELECTRICS

575. Conductors and Insulators. In conductors the slightest external electric force causes electrons to pass continuously from atom to atom through the body, thus constituting a flow or current of electricity. In this motion the electrons are constantly checked by their impacts against atoms of matter or other electrons and in this way they are retarded by a sort of

frictional resistance, just as shot are retarded in moving through a mass of molasses; but there is nothing like an *elastic* resistance to make them spring back when the displacing force is removed.

In insulators, on the other hand, if electric force is applied, there is a certain yielding or displacement of the electrons. If the force is increased the electrons are displaced more, but there is no continuous flow as in a conductor, and as soon as the external force is removed they spring back to their original positions, behaving as shot would if imbedded in a mass of rubber.

Now suppose that the process of charging two conductors *A* and *B* by an electric machine consists in forcing some electrons out of *A* into *B*, thereby making *A* positive and *B* negative. This will cause a crowding outward of the electrons in the dielectric immediately surrounding *B*, while those around *A* will be displaced inward to make up for its deficiency, and so in all the dielectric surrounding *A* and *B* the electrons will everywhere be displaced in the opposite direction to the arrows which indicate the positive direction of the lines of force.
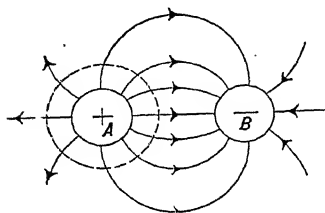
Fig. 311

But since in a dielectric electrons are not free to move, their displacement at any point is only through a very small distance and is opposed by the internal electric forces between the positive and negative elementary charges in the dielectric which urge all the electrons back toward their original unstrained positions. There is therefore produced a back pressure on the electrons in *B* and a negative pressure on that in *A*, so that if *A* and *B* are now connected by a conducting wire there will be a flow of electrons from *B* to *A*, until the displaced electrons in the dielectric have sprung back into their original positions. *The discharge is thus conceived as forced from B to A by the springing back of the electrons in the dielectric in consequence of the internal electric forces in the dielectric.*

This difference in pressure between *A* and *B* due to the reaction of the displaced dielectric is the difference between their potentials. Suppose that after *A* and *B* are charged they are moved nearer together. The strain will now take place through a less thickness of dielectric and the difference in pressure between *A* and *B* will accordingly be less. The work required to produce a given charge will therefore be less when they are nearer together; that is, the energy of the charge will be less. They will therefore tend to move together; that is, there is an *attraction* between *A* and *B*. For if they are moved apart they will have more energy, but they can only get this additional energy from the work done in separating them; therefore there must be a force opposing the separation or a force of attraction.

576. Tubes of Force. The electric field may be conceived as divided up into tubes by means of surfaces in the direction of the lines of force. (Compare § 511.) These tubes of force will always have *at one end a positive charge and at the other an equal negative charge.*

On the electron theory there will be as many electrons displaced inward across one end of a tube of force as will be displaced outward across the other end.

577. Induction as Explained on Electron Theory. Suppose A and B are conductors near each other (Fig. 312) and having no charge at first. Let a negative charge be given to A.

In doing this we may suppose that electrons are transferred from the ground so that the walls of the surrounding room become positive. Tubes of force in which electrons are displaced outward will extend outward from A in every direction toward the walls of the room. But since B is a conductor there is no force resisting the displacement of the electrons through it,
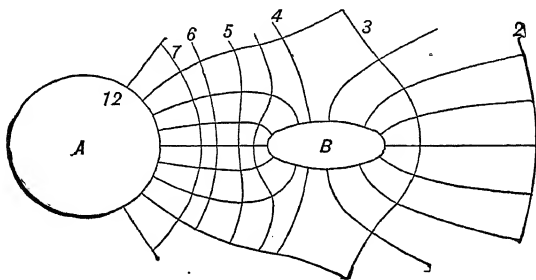
FIG. 312. Induction

whereas in every other direction there is the active elastic resistance of the dielectric to be overcome. Displacement can therefore take place more readily on the side of A toward B than in other directions and a number of tubes of displacement will terminate on B, and the electrons where the tubes terminate will be displaced toward B's interior, while on the farther side of B there will be an equal outward displacement of electrons. These constitute equal positive and negative charges respectively.

578. Why an Electrostatic Charge Appears Only on the Surface of a Conductor. If a charge is given to a hollow conductor (Fig. 313) all parts of the metal shell come to the same potential and there is a displacement of electrons in the dielectric surrounding it, and this displacement is either away from the conductor or toward it depending on the kind of charge given to the conductor.

But the dielectric A in the interior *is entirely surrounded by the conductor and is therefore pressed upon equally in every direction*, consequently there can be no displacement of its electrons. The pressure or potential in the interior is, however, everywhere the same as that in the conductor which surrounds it. If a small conductor B is touched to the interior of the shell it comes to the same potential as the shell, but the electrons in the dielectric around it, being equally urged in every direction, are not displaced and accordingly B neither gains nor loses electrons; that is, it does not receive a charge.

It is much as though a bottle with flexible rubber walls was filled with

water and then put inside of a vessel full of water under considerable pressure. If the stopcock is then opened no water will flow either into the bottle or out of it. For the pressure is the same inside of the bottle as it is outside. If the stopcock is closed and the bottle is removed from the region of pressure it will be found neither to have gained nor lost charge.

So it is also with the conductor B; when inside of A it is at the same potential as A, but it does not receive any charge because no displacement in the dielectric at its surface is possible, and so when removed from A its potential changes to that of the region where it is placed, but it shows no trace of charge.

579. A Case of Induction. Suppose, however, that the conductor B, while inside the charged body A and insulated from it, is connected with
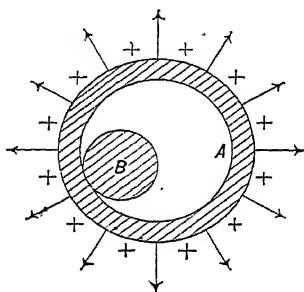
FIG. 313.   Hollow conductor   FIG. 314
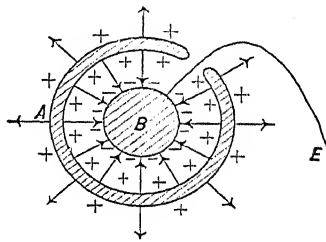
the earth by a wire as shown in figure 314. The electrons in the dielectric will then be displaced not only in the dielectric between A and the walls of the room, but also between A and B, and if A is positive they are displaced toward A and away from B causing a corresponding flow of electrons into B from the earth.

If B is now insulated and removed from the interior of A it will be found negatively charged, for the displacement in the dielectric around it cannot disappear until the electrons that flowed into B are permitted to escape.

Since the outward displacement on B must be equal to the inward displacement over the inner surface of A, the charge on B must be equal and opposite to that on the interior of A; this has previously been shown to be the case in Faraday's ice-pail experiment (§ 556).

CONDENSERS AND CAPACITY

580. Condenser Experiment. Take a tin plate, mount it bottom upward on an insulating stand and connect it with a gold-leaf electroscope. Cover the plate with a sheet of glass and then give it a sufficient charge to cause the gold leaves to diverge

strongly. Now take another tin plate, connected to earth by the hand or by a wire, and lower it upon the glass. The gold leaves will be observed to come together as the plates approach each other, showing that the potential of the charged conductor is diminished by the approach of the grounded conducting plate. The closer the two plates are brought together, the greater will be the decrease in potential. On removing the upper plate the leaves diverge as at first, showing that there has been no loss of charge.

In this experiment evidently the *capacity* of the first plate has been increased by the proximity of the second uninsulated plate. Such a combination is known as a *condenser*, because it can take a large charge at a small potential.

The decrease in potential is due to the presence of an induced charge on the second plate opposite in kind to that on the first.

581. Leyden Jar. The earliest form of condenser was devised in 1746 by Musschenbroek, of Leyden. That experimenter, in attempting to charge a glass of water with electricity, held the glass in his hand while one pole of the electrical machine was connected with the water through a nail resting in the glass. After charging it well, the knuckle of the other hand was touched to the nail and a smart electric shock was obtained.

FIG. 315. Discharge of Leyden jar

Further experimentation showed that all that was necessary was that there should be two conducting coatings separated by the glass.

Accordingly, a *Leyden jar*, as it is called, is made by coating a glass jar or bottle inside and outside with tinfoil for about two-thirds the height of the jar. Connection is made with the inner coating through a metal rod terminating in a knob.

To charge such a jar one coating must be connected to one pole of the electrical machine while the other coating is connected to the other pole of the machine either directly or through the earth, so that one coating receives electrons while the other loses them, thus producing opposite charges upon the two coatings.

The jar is discharged by connecting the knob and outer coating by a conductor.

If the outer coating is touched with one hand while the knuckle of the other hand is brought to the knob, the jar is discharged through the body and the sensation of shock is experienced. Slight shocks are felt in the hand and arms while stronger shocks are felt in the body.

A Leyden jar is said to have greater *capacity* for charge than an ordinary insulated conductor, because it will receive a much greater charge from a given electrical machine than will be taken by the simple conductor.

582. Condensers. Any contrivance in which two conductors are separated by a thin dielectric which has sufficient dielectric strength to prevent the discharge of electrons between them, is a condenser and has the same properties as a Leyden jar.

A convenient form of condenser, due to Franklin, may be made by coating the opposite sides of a plate of glass with tin-foil, which, however, must not reach too near the edges of the plate.

583. Insulated Leyden Jar. *If a Leyden jar has either of its coatings insulated, no more charge can be given to the other coating than to a simple metal conductor of the same shape and size.* If a Leyden jar is charged and then placed on an insulating stand, it cannot be discharged by simply touching the knob. But if the finger is touched first to the knob and then withdrawn and touched to the outer coating and so on *alternately* a small spark will be obtained every time and the jar may thus be very slowly discharged.

584. Explanation of Action of a Condenser. One way of explaining the action of a condenser is as follows: The plate A, say, receives electrons from the electrical machine. This negative charge acts inductively through the glass dielectric, repelling the electrons in the other plate so that some of them flow off into the earth. This leaves this plate positive so that its attraction enables the machine to give the plate A a much larger negative charge. If the plate B were not connected to the earth, its electrons could not escape even though repelled by the plate A. These trapped electrons, by their repulsion would prevent A from receiving any more electrons than if the plate B were not there.

But it is better to look at the action in the following way,

from the standpoint of the displacement theory. Let A (Fig. 317) represent the inner coating of a Leyden jar which is stripped of its outer coating. Connect A to the positive pole of an electrical machine and connect the negative pole to the floor or walls of the room. The conductor A will become positively charged and an equal and opposite negative charge will be found on the walls of the room. The tubes of force or displacement extend from A to the walls,* but the difference of potential produced by the machine *can cause only a small strain or displacement of the electrons in so thick a dielectric, and, therefore, only a small charge will be given to A.* But if the outer coating is now put upon the jar and connected with the negative pole of the machine, as in figure 318, the whole strain will take place in the thin layer of glass and so a great displacement of electrons in the glass will take place involving a large flow of electrons into one coating and out of the other, leaving the jar strongly charged.

FIG. 316. Charging a condenser

If the outer coating were *insulated* and the negative pole of the charging machine were connected to the walls of the room,
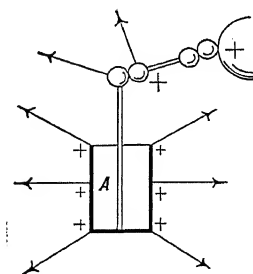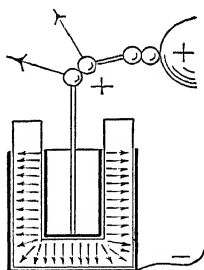
FIG. 317   FIG. 318

no displacement could take place through the glass toward the outer coating without an equal displacement taking place

---

* It should be kept in mind that the convention adopted for the direction of tubes of force shows them pointing in the opposite direction to that of the actual electron displacement in the dielectric.

outward from the outer coating toward the walls, so that the jar would take no greater charge than if there were no outer coating.

585. Capacity of a Conductor. From the above discussion it will be clear that *the charging of every conductor is analogous to the charging of a Leyden jar;* the conductor itself corresponds to the inner coating of the jar; while the surrounding walls or conductors on which the tubes of force terminate correspond to the outer coating; *it differs from a Leyden jar only in the greater thickness of the dielectric.*

586. Capacity. The larger the charge given to a Leyden jar or condenser, the greater is the difference of potential between its coatings, so that we have

$$Q = VC$$

where Q is the charge, V is the difference of potential between the coatings of the jar, and C is a constant called the *capacity* of the jar.

When $V = 1$, $C = Q$, and, therefore, *the capacity of a condenser is numerically equal to the quantity of charge required to make unit difference of potential between its coatings.*

Capacity depends on that area S of one surface which is opposed by the other and varies inversely as the thickness of the dielectric d which separates them. It may be computed from the formula

$$C = \frac{SK}{4\pi d}$$

where K is a constant which depends on the nature of the dielectric and is called its *specific inductive capacity* or *dielectric constant.*

The derivation of this formula is given in § 596.

CAUTION: The student is warned against thinking that the capacity of a condenser is the greatest charge which it can hold. Capacity is not a charge or quantity of electricity, but a constant of the condenser which does not change with different charges. It is the *ratio* of charge to difference of potential between the coatings, and remains the same for all charges. The maximum possible charge of a condenser depends upon its insulation and the strength of the dielectric between its coatings to resist disruptive discharge.

Some Specific Inductive Capacities, or Dielectric Constants

Hard rubber . 2.5	Paraffin.... 2.0	Air (normal pressure). 1.00059
Glass..... 6 to 8	Turpentine . 2.2	Carbon dioxide....... 1.00090
Mica........ 8.0	Petroleum.. 3.1	Hydrogen........... 1.00028

587. Hydraulic Analogy. It is instructive to consider the following hydraulic model of a condenser. A metal box is divided into two parts A and B by a partition of thick sheet rubber. Each side is provided with a tubular opening controlled by a stopcock, the whole is then filled with water and immersed in a pond. While the stopcocks are open the pressure is the same in A and B and the rubber is not strained. It is like a Leyden jar uninsulated and discharged. Now attach a pump to B and force water in while the stopcock of A is left open or, what amounts to the same thing, connect the pump both to A and B so that it pumps water out of A and into B. The rubber will be strained as shown in the figure, the side A will be at the pressure of the pond which may be called zero, while the other side is at a

higher pressure p. This difference in pressure p between the two sides is due to the strain of the rubber. If A and B are now connected by a pipe and the stopcocks are opened there will be a flow from one side to the other as the rubber springs back into the unstrained condition.
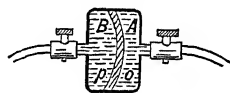
Fig. 319. Hydraulic model of Leyden jar

So when a Leyden jar is discharged electrons may be thought of as forced from one coating to the other by the springing back of the displaced electrons in the dielectric.

If the rubber diaphragm were thicker more difference in pressure would be required to force in a given charge. So in a Leyden jar, the thicker the dielectric the greater the difference of potential between its coatings when it has a given charge.

If the diaphragm were made of a substance that was more *yielding* than rubber, it would correspond to a dielectric of *greater specific inductive capacity;* for a given pressure would then force in a greater charge.

Also suppose the stopcock A is closed and the pump connected to B, pressure will be produced in B and perhaps a slight amount of water forced in due to the elastic yielding of the box itself, but the rubber diaphragm will not be appreciably strained and the pressure will be the same on both sides. This is the case of trying to charge an insulated jar. The stiff and but slightly yielding walls of the box represent the insulating dielectric that surrounds the Leyden jar and extends to the walls of the room, while the rubber diaphragm represents the thin glass dielectric between the coatings of the jar.

Remembering that the dielectric surrounding the jar is slightly yielding will enable the student to explain the succession of small sparks obtained from the insulated jar as described in § 583.

588. Energy of Charge. When we begin to charge a Leyden jar or condenser the two coatings are at the same potential and

therefore no work is required to transfer the first few electrons from one coating to the other. But as the charging goes on the difference of potential between the two coatings increases and more work is required to transfer a given number of electrons.

Suppose the final potential to which the jar is charged is V, and suppose that in charging it Q units of electricity are transferred from one coating to the other, giving one a charge $+Q$ and the other a charge $-Q$. If during this transfer the difference of potential between the coatings were to remain constant and equal to V the work done in charging would be QV ergs. But since the difference of potential is zero at the start and increases in proportion to the charge, *the average potential during charging is $\frac{1}{2}V$ and the work actually done in charging is $\frac{1}{2}QV$, which is therefore the energy of the charge.*

The case is analogous to the filling of a cylindrical water tower, the pressure is zero when the tower is empty, and increases
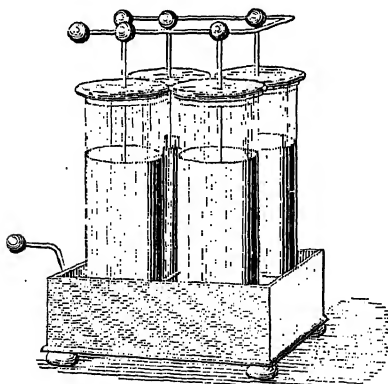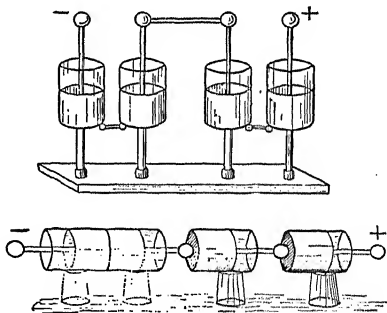
FIG. 320   FIG. 321

as the water rises until the final pressure p is reached. The work done is, therefore, $\frac{1}{2}pv$ where v is the total volume of water pumped in.

The energy of the condenser exists as electrical strain in the dielectric.

589. Leyden Battery. The Leyden jars in the combination shown in figure 320 have their inner coatings connected together and are mounted in a box lined with tinfoil by which their outer

coatings are also joined.   Such an arrangement is known as a *Leyden battery*, the jars are also said to be connected in *parallel* or *multiple*, and the combination is equivalent to a single large jar having a *capacity equal to the sum of the capacities of the separate jars.*

590. Leyden Jars Connected in Cascade or Series. In each of the two arrangements shown in figure 321 four jars on in-
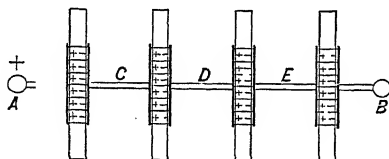
Fɪɢ. 322.   Condensers in series

sulating stands are connected in such a way that if the discharge were to burst through the glass of the jars it would have to pierce all four jars to pass from one end to the other, as four layers of glass intervene between the terminal conductors.   In such a case the jars are said to be joined in cascade or in series.   The diagram (Fig. 322) shows that state of electrification, the regions between the charged plates representing the layers of glass.

Four similar jars joined in this way are like a single jar having a dielectric four times as thick, and the capacity of the combination is one-fourth that of a single jar.

This case is well illustrated by the hydrostatic analogue (Fig. 323) in which four models such as are described in § 587 are connected in series. Clearly when water is pumped in at A and out at B the rubber diaphragms

Fɪɢ. 323

are all strained and an equal quantity of water is displaced from each into the next succeeding, thus representing the equality of the charges in each. The pressures represent the potentials.   Evidently the pressure p_1 is greater than p_0, and p_4 is the greatest of all, and to force in a given quantity of water four times as much pressure must be used as to force it into a single one of the cells.

The chief practical advantage of the cascade arrangement is that it has *great dielectric strength* and sparks do not easily burst through the glass; for

this reason the small jars used on induction electrical machines are usually connected two in series, one being connected to one pole of the machine and one to the other, while their outer coatings are joined by a wire.

When Leyden jars of different capacities C_1, C_2, C_3, are joined in series, the capacity C of the combination is found from the relation

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}$$

To prove this formula let Q represent the charge, which will be the same for each jar when they are charged in series. The potentials of the jars will be

$$V_1 = \frac{Q}{C_1}, \quad V_2 = \frac{Q}{C_2}, \quad V_3 = \frac{Q}{C_3}.$$

The total difference of potential between the end coatings of the series will therefore be

$$V = V_1 + V_2 + V_3 = Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right)$$

and if C is the capacity of the combination we have

$$V = \frac{Q}{C}; \text{ therefore, } \frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}$$

PROBLEMS

1. A Leyden jar 14 cms. in diameter and made of glass 3 mm. thick is coated on the bottom and sides up to a height of 20 cms. What is its capacity and what charge is required to bring it to potential 30? Take dielectric constant of glass = 6.

2. Two Leyden jars, one of capacity 300 and charged to potential 20, the other brought to potential 30 by a charge of 7200 units, are connected together in parallel, positive coating being connected to positive and negative to negative. Find the resulting potential and the charge in each jar after being connected.

3. If in the preceding problem the positive coating of each jar is connected to the negative of the other, what will be the resulting difference in potential and charge in each jar?

4. A jar of capacity 1000 is charged to potential 50; find the heat developed in gram-calories when it is discharged through a long fine wire.

5. Three jars each of capacity 500 are charged each to potential 12 and then joined in series and finally discharged by connecting the end coatings

of the combination. Find the difference of potential between the end coatings and the quantity of charge that passes through the discharge, and thence calculate the energy expended in the discharge.

6. Let the three Leyden jars of the previous problem be joined in parallel and then discharged. Find the difference of potential between coatings, and the quantity of discharge, and thence determine the energy of the discharge.

7. A Leyden jar of capacity 500 is joined in series with another of capacity 200, and the combination is given a charge 3000. Find the difference of potential between the coatings in each jar. Thence find the difference of potential between the end coatings of the combination. What is the capacity of a single jar which when given the charge 3000 would have the same difference of potential as the combination?

8. A Leyden jar of capacity 600 is joined in series with one of capacity 400. Find the capacity of the combination.

9. Two Leyden jars of capacities C_1 and C_2 are joined in series and given a charge Q. Find the capacity of a single jar equivalent to the combination, by following the method of problem 7.

10. A Leyden jar of capacity 800 is joined in series with another of capacity 200, and the combination charged to potential 20. Find the charge in each jar and the difference of potential between the coatings of each jar.

CALCULATION OF POTENTIAL AND CAPACITY

591. Potential at a Point. Up to this point we have thought of electrical potential simply as a certain condition which determines the flow of electricity; and we have shown that the difference of potential between two conductors may be measured by the work done in transferring unit charge from one to the other (§ 571).

But potential is not a property of conductors only. When a little charge is brought up to any point whatever in space, work must in general be expended in bringing it to that point, on account of the attractions or repulsions of neighboring charges; and this work, per unit charge, is used as the measure of the potential at that point.

Definition. *The potential at any point is measured by the work done against electrostatic forces in bringing a unit positive charge up to that point from an infinite distance.*

This work may be calculated as follows:

Suppose there is a charge of q units of electricity at A

(Fig. 324), and it is required to find the work done in carrying unit charge from C to B in the same straight line with A when air is the medium between the charges.

Conceive the distance BC divided into n small parts at the points a_1, a_2, a_3, etc., and let r be the distance from A to B and r_1 the distance from A to a_1, etc. Then the force with which q repels unit charge at B in air is $\dfrac{q}{r^2}$ (§ 540), while the force at a_1 is $\dfrac{q}{r_1^2}$.
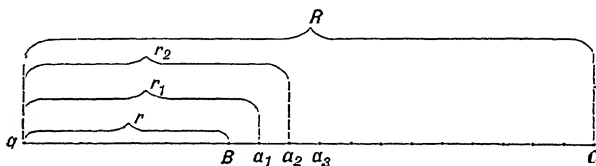
FIG. 324

To get the work w_1 done when unit charge is moved from a_1 to B, the average force must be multiplied by the distance from B to a_1 which is $r_1 - r$.

The geometrical mean of $\dfrac{q}{r^2}$ and $\dfrac{q}{r_1^2}$ is $\dfrac{q}{rr_1}$; and this may be taken as the average force between a_1 and B if these points are close together. The work done in this element of the distance will therefore be

$$w_1 = \frac{q}{rr_1}(r_1 - r) = \frac{q}{r} - \frac{q}{r_1}$$

and similarly the work done when unit charge is moved from a_2 to a_1 is

$$w_2 = \frac{q}{r_1} - \frac{q}{r_2};$$

so also

$$w_3 = \frac{q}{r_2} - \frac{q}{r_3};$$

and finally

$$w_n = \frac{q}{r_{n-1}} - \frac{q}{R}.$$

Adding, we find

$$w_1 + w_2 +, \text{ etc., } + w_n = \frac{q}{r} - \frac{q}{R}$$

where $w_1 + w_2 +$, etc., $+ w_n$ is the whole work done in moving the unit charge from C to B. In the final result all intermediate

terms have disappeared, the result is therefore the same however great the number of parts into which CB may be divided; it is therefore clear that no error was introduced by taking the geometrical mean of the forces at B and a_1 as the average between those points.

It may be shown that the work will be the same along any path whatever between B and C, even though these points may not lie in the same direction from A.

Thus the work done in carrying unit positive charge from C to B (Fig. 325) against the repulsive force of a charge q at A is

$$\frac{q}{R}.$$
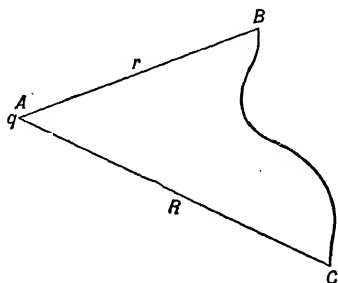
Now if the point C is at an infinite distance from A then $\dfrac{R}{q}$

FIG. 325

$= 0$, and the work done against the repulsion of q, in bringing a unit charge up to B from an infinite distance in air or vacuum, is simply $\dfrac{q}{r}$, and this, by definition, is the potential at B due to the charge q. Representing this potential by V we have,

$$V = \frac{q}{r}.$$

If there are a number of charges q_1, q_2, q_3, etc., at distances r_1, r_2, r_3, respectively, from the point B and in any directions whatever, the potential at that point becomes, when air is the medium,

$$V = \frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} +, \text{ etc.}$$

since *the potential at a point depends only on its distance from charges and not on their directions*. The signs of the terms depend on whether the charges are positive or negative.

In any other medium than air the potential V may be computed from the formula

$$V = \frac{1}{K}\left(\frac{q_1}{r_1} + \frac{q_2}{r_2} + \frac{q_3}{r_3} +, \text{ etc.}\right)$$

where K is the specific inductive capacity of the medium (§ 540).

592. Zero Potential. According to the definition just given, those points are at zero potential which are at an infinite distance from all electrified bodies. But the earth's potential has also been defined (§ 567) as zero potential. These two definitions are inconsistent if the earth has a charge, and there are reasons for thinking that it has.

But any charge which the earth may have will change the potential of the earth and of all bodies in our laboratory rooms by the same amount, so that *differences of potential will be unchanged*, and it is only differences of potential that are measured by our instruments.

In discussing problems that involve the electrical state of the heavenly bodies or of regions remote from the earth, of course it will not do to assume that the earth potential is zero. The zero must then be taken as defined in the preceding paragraph.

593. Equipotential Surfaces. Suppose there is a charge of 12 units at A (Fig. 326) which is not near any other charged body. Then the potential due to A may be calculated from the formula

$$V = \frac{q}{r}$$

where $q = 12$. At 1 cm. from A in any direction the potential will be 12. The sphere of radius *one* having A as center is therefore an equipotential surface of potential 12. The sphere whose radius is 2 cms. is the surface of potential 6, the surface of potential *one* would have a radius of 12 cms., while zero potential would be at an infinitely great distance. If the charged point A is inside of a room the surface of the room will be at zero potential, for there will be an induced negative charge at each point of the surface sufficient to counteract the action of the charge A.

The figure shows the position of the successive equipotential surfaces, differing by unity, from 2 to 12. It will be noticed that they are closer together the nearer they are to A. The same amount of work must be done to move a unit charge from the surface 2 to 3, as from 3 to 4 or 11 to 12; in each case one erg of work is done. *But the shorter the distance in which a given amount of work is done the greater the force that must be exerted, hence*

the surfaces are closer together near A where the electric force is greater.

No work at all is done when an electric charge is moved along an equipotential surface, hence *at every point the direction of*
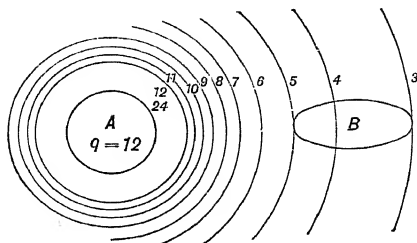
FIG. 326.   Equipotential surface due to a charge 12 at *A*

the resultant force must be at right angles to the equipotential surfaces.

Lines of force, or lines which at each point have the direction of the resultant force, must therefore cut equipotential surfaces at right angles, and in the above case are a set of radial straight lines.

594. Induction from the Point of View of Equipotential Surfaces. In figure 326 notice that the region *B* surrounded by
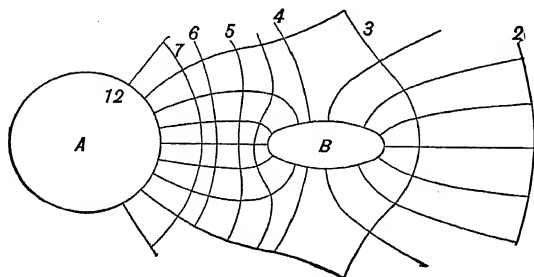
FIG. 327.   Equipotential surfaces where *B* is a conductor

the elliptical line reaches from a point where the potential is 3 to where the potential is 5.   If *B* is a non-conductor this distribution of potential is possible, *but if B is a conductor flow must take place until it is all at the same potential.*   The left-hand end will receive a negative charge which will lower its potential, while the right-

hand end will have its potential raised by a positive charge till all parts come to some potential intermediate between 3 and 5. The lines of force and equipotential surfaces in this case are shown in figure 327. In this diagram the conductor is supposed to come to potential 4, all other equipotential surfaces are bent outward or inward away from B. Some lines of force from A terminate on the negatively charged left end of B, while lines of force go out from the positively charged end of B to the right. This case is analogous to the formation of a level spot or pond on the side of a mountain. The ground must be cut away on the side toward the mountain, and built up on the outside.
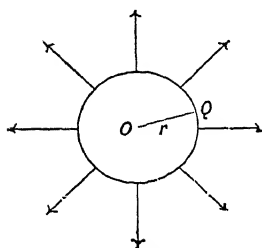
FIG. 328. Isolated sphere

595. Capacity of an Isolated Sphere. Suppose an insulated sphere in the center of a large room. If it has a positive charge Q, its lines of force terminate on an equal quantity of negative electricity induced on the walls of the room. To find the potential at the center of the sphere we have the formula,

$$V = \frac{q_1}{r_1} + \frac{q_2}{r_2} + , \text{ etc.} \qquad (\S 591)$$

In the present case the only charge which is near enough to produce any appreciable effect at O is the charge $+Q$. Although this charge is distributed over the sphere, it is all at the same distance r from O. Therefore the potential at the center of the sphere is

$$V = \frac{Q}{r}.$$

But in case of a charged conductor all parts of it, inside and outside, are at the same potential, the sphere is, therefore, all at the potential V of its center.

But by § 586

$$Q = VC,$$

therefore

$$C = r$$

or the capacity in electrostatic units of an isolated sphere surrounded by air is numerically equal to its radius.

If the medium surrounding the sphere has specific inductive capacity K, its capacity becomes (§ 586)

$$C = Kr.$$

596. Capacity of a Condenser Made of Two Concentric Spheres. Suppose we have a condenser such as shown in figure 329, consisting of two concentric metal spheres with air between them. Let r_1 be the outer radius of the inner sphere and r_2 be the inner radius of the outer sphere. If a charge $+Q$ is given to the inner sphere, an induced charge $-Q$ will be found on the outer sphere. If the outer sphere is connected to earth it comes to zero potential and all charge disappears from its *outer* surface.
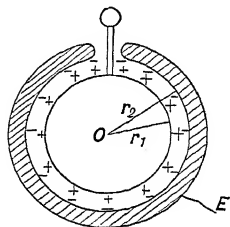
Fig. 329. Spherical condenser

The potential at O the center of the small sphere is therefore

$$V = +\frac{Q}{r_1} - \frac{Q}{r_2}$$

since the charge $+Q$ is at a distance r_1 from the center, and the charge $-Q$ is at a distance r_2 from the center.

But the potential everywhere inside of a closed conductor is the same as at its surface. Hence the potential of the inner sphere is

$$V = +\frac{Q}{r_1} - \frac{Q}{r_2} = Q\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

and since the outer sphere is at zero potential, V is the difference of potential between the two.

But by the definition of capacity $Q = CV$, therefore

$$C = \frac{1}{\dfrac{1}{r_1} - \dfrac{1}{r_2}} = \frac{r_1 r_2}{r_2 - r_1}$$

If the medium between the spheres has a specific inductive capacity K, the capacity of the condenser will be

$$C = \frac{K r_1 r_2}{r_2 - r_1}$$

If the spheres are close together we may write $r_1r_2 = r^2$ and $r_2 - r_1 = d$ where r is the mean radius of the spheres and d is the thickness of the space between them.

Then

$$C = \frac{Kr^2}{d} = \frac{K4\pi r^2}{4\pi d}$$

but $4\pi r^2$ is the area of surface of a sphere of radius r; therefore

$$C = \frac{KS}{4\pi d}.$$

In this form the formula can be used for any condensers where the two surfaces are close together, as in a Leyden jar or in a condenser made of two flat parallel plates.

PROBLEMS

1. How much work must be done to carry a unit positive charge from a point 1 meter distant from a charge $+100$ to a point 2 cms. from it?

2. What is the potential at a point half-way between two equal spherical conductors having charges $+100$ and -100, respectively?

3. What is the potential at one corner of a rectangle which measures 40×30 cms. when there is a charge -300 at the diagonally opposite corner and $+120$ at each of the adjacent ones?

4. A spherical conductor 10 cms. in diameter has a charge of $+200$ units and a small body having an equal plus charge is situated 1 meter from the center of the sphere. What is the potential at the center of the sphere? What is the potential at its surface? Is the charge distributed uniformly over the sphere?

5. How much work would be done in moving the small charged body of the preceding question up to 50 cms. from the center of the sphere?

ELECTRIC DISCHARGE

597. Electric Discharge through Air at Ordinary Pressure. Three forms of discharge through air are recognized at ordinary pressures, the *electric spark* or *disruptive discharge, brush discharge*, and *glow discharge*.

In the ordinary spark discharge there is a rush of electricity accompanied by a flash of light and by heat and sound.

The *brush discharge* is seen when in a darkened room the hand is brought near the positive conductor of a highly active electrical machine. If it is not held near enough for the spark discharge a luminous brush, like a little tree with branches of light ramifying from a short stem, extends out toward the hand from some point on the positively charged conductor. It seems to be caused by an almost continuous succession of extremely small discharges.

Sometimes in the dark when an electric machine is highly excited, but when the conductors are separated too far for sparks to pass, a faint velvety glow of violet light known as the *glow discharge* is seen on the knob of the negative conductor.

All these forms of spark discharge are first started by a portion of the molecules being broken up each into a positive ion and a negative ion or electron (§ 769). This is caused by the electric force acting on the few positive and negative ions, which are always present in air, so that they collide violently with other molecules and knock off some of their electrons. These new ions are also acted upon by the electric field, the positive ions being driven in one direction and the negative ions or electrons being driven in the opposite direction. These in turn produce more ions by collision with neutral molecules and, aided by a supply of electrons which are apparently liberated at the negative electrode by the arrival there of positive ions, the current almost instantly builds up to a large value and is made evident as the spark discharge. The spark is of very short duration since, owing to the large current, the conductors are almost instantly discharged.

If the electric force producing the ions is fairly uniform in intensity, as it is between the knobs of an electric machine, when the electric force reaches a certain value, this *ionization* suddenly builds up all the way across between the knobs which starts a spark discharge. On the other hand, for the case of brush discharge the electric force is intense enough to produce active ionization only at sharp points and projections of a charged body instead of along an entire path, so that only a slow leakage of charge can take place.

598. Oscillatory Discharge. When a spring is bent and let fly it oscillates back and forth, coming to rest when its energy is

finally spent in heat, sound, and air waves. So when a charged Leyden jar is discharged through a circuit of small resistance the energy of the charge cannot be dissipated in the first rush and consequently there is a back-and-forth rush of electrons from one coating to the other until the energy is finally spent in sound,
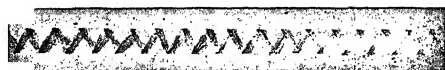
FIG. 330.   Oscillatory discharge

heat, light, and electric waves. This is known as the oscillatory discharge. If there is sufficient resistance in the discharge circuit there is no oscillation, just as a pendulum hung in molasses will sink to its lowest position without oscillation.

The oscillatory discharge was examined by Feddersen in 1863 by means of a rapidly rotating mirror. Seen in this way, each discharge showed as a group of sparks at regular intervals and rapidly dying out, as shown in figure 330; see also § 811.

599. Mechanical and Heating Effects of Disruptive Discharge. When the discharge takes place through a sheet of glass it is pulverized at the point of discharge. Pasteboard is perforated by the discharge, the edges of the hole being raised in a burr on each side as if by the sudden expansion and bursting out of the contained air or moisture. When trees are struck by lightning they are apt to be splintered, large slivers being flung violently out sidewise, perhaps due to sudden vaporization of moisture. When a living tree is struck the discharge usually takes the sap layer and frequently follows the grain. A glass tube having a fine bore filled with water and with a wire thrust a short distance in each end may be burst by the discharge of a Leyden jar.

The electric discharge is accompanied by heat. Ether and bisulphide of carbon are readily ignited by it. Buildings containing inflammable material are occasionally set on fire when struck by lightning. A mixture of one volume of oxygen with two volumes of hydrogen explodes with violence if even a minute electric spark passes in it.

A little gunpowder placed between the ends of two wires through which a discharge is sent will usually be scattered unless the discharge is retarded by causing it to pass through a wet

string or other poor conductor, in which case the powder may be ignited.

Narrow strips of gold foil, 1 or 2 mm. in width, gummed to a sheet of paper so that they form a conducting strip, may be deflagrated or volatilized by the discharge of a Leyden battery. The purple stain which is left is wider than the gold-foil strips and is streaked at right angles to its length as though the volatilized metal had been driven violently out from the path of discharge.

600. Lightning. The resemblance between lightning flashes and electric sparks was early noticed. Franklin, in 1752, performed the celebrated experiment of obtaining electric charges by means of a kite as a thunderstorm was approaching. The kite was provided with metal points and the linen kite cord was a fairly good conductor when wet. To the lower end of the kite cord was fastened a metal key to which a silk cord was attached which was held in the hand and acted as an insulator. Sparks were obtained from the key and Leyden jars were charged, and the familiar phenomena of electric charges were observed.

The so-called *globe lightning*, described by different observers as a ball of fire slowly moving along and then suddenly exploding with terrific violence, has never been imitated by any electrical discharges obtained in the laboratory and is so different from the ordinary phenomena of discharge that many physicists consider such observations illusory and due to a subjective effect of the discharge on the eye of the observer.

601. Atmospheric Electricity. The electrical separation in thunderstorms according to the theory of Simpson, is due to the disruption of rain drops in the uprushing current of air; for laboratory experiments show that when a drop is broken up by falling on a vertical jet of air the resulting drops are positively charged while the current of air carries off negative charge. Rain from the lower part of the cloud will carry down positive charge while rain from higher regions of condensation will be negatively charged.

Another circumstance that very possibly plays a part in the development of thunderstorms is that condensation of moisture in the atmosphere takes place more easily around negative nuclei or electrons than it does around positive nuclei, and the

fall of such drops to the earth will give it a negative charge. In fair weather the earth is usually negative, the potential being higher at points above the earth's surface, increasing at the rate of from 75 to 150 volts per meter above level ground, while in thunderstorms the atmospheric potential fluctuates greatly and may even be negative to the earth.

602. Lightning Rods. It was shown (§ 553) that when an electroscope was surrounded by a conducting surface or even enclosed in a wire cage it was screened from outside electrical disturbances, and this suggests how buildings should be protected.

Buildings with metal outer sheathing need no other protection, though care should be taken that the metal walls are at least as well connected to damp earth as the gas and water pipes within.

Wooden structures should have low metal points on the chimneys and gables and other projecting portions, these points should be connected together by heavy wires or other conductors which run down the main corners of the building to the ground. At or near the ground it is well to have them connected together by a wire passing entirely around the building, and at two points on opposite sides of the building good ground connections should be made by connecting to pipes driven down to water or to a metal plate bedded in coke in damp earth.

Insulation from the building is not needed, metal roofs and gutters and rain-water pipes should be connected together and may serve for lightning conductors if given good ground connections. Ordinary heavy galvanized iron telegraph wire will serve well for the conductor or, still better, a flat ribbon of sheet copper.

603. Piezo Electricity. Certain crystals, such as tourmaline, quartz and notably Rochelle salt become electrically charged when they are mechanically stressed. If the applied stress be reversed, that is, if tension is changed to compression, or if a right-hand twist is changed to a left-hand twist, the sign of the charge which appears on the crystal becomes reversed, also.

Rochelle salt is many times more active than any other crystal known. A force of 10 kilograms applied to such a crystal in the right relation to its axes may produce charges equal to more than 100 electrostatic units.

The piezoelectric effect, as it is sometimes called, is reversible,

that is, when the crystal is charged first, it becomes distorted as though a stress were applied to it, this effect being the reciprocal of the production of electrification by applying stress.

Microphones for use in radiotelephony (§ 824) have been designed which make use of piezo electricity. The alternating pressure of the sound waves of the voice act upon a suitably supported crystal of Rochelle salt and produce electric charges upon it which oscillate exactly according to the sound waves. These electric charges are amplified by three electrode vacuum tubes (§ 781) and transmitted in the usual way. These microphones have the advantage of responding almost equally to all frequencies, but the disadvantage of varying in sensitivity with temperature variation.

The reciprocal effect may be strikingly demonstrated by applying alternating electric charges of high frequency from an energetic vacuum tube oscillator to the sides of a small slab cut from a quartz crystal, perpendicular to the crystal axis. An alternating expansion and contraction of the quartz is produced. When the oscillator frequency is tuned to the natural mechanical vibration period of the piece of quartz crystal, the vibration of the piece of quartz may become so violent as to cause it to burst into fragments due to the mechanical stress developed.

ELECTRIC CURRENTS

OR

ELECTRODYNAMICS

THE ELECTRIC CURRENT AND VOLTAIC CELL

604. The Electric Current. When a Leyden jar is discharged or when a series of sparks from an electrical machine pass through a conductor, in fact whenever a charge is communicated from one point to another, there is what is called a flow of electricity, or an *electric current*.

The current is said to flow from the positive to the negative conductor. This is a convention; for vitreous electrification was called positive, and resinous was called negative, long before there was any idea of the direction of flow. It is now known that *an electric current in a conductor consists of a flow or transfer of elementary negative charges or electrons from the negative to the positive conductor*, thus the actual direction of flow is exactly opposite to the ordinary convention.

In all the cases hitherto considered the flow has been so transitory as to be almost instantaneous. We now come to a series of discoveries which made possible the production of currents of electricity lasting for a considerable time.

605. Galvani's Discovery. In 1786, Galvani, professor of anatomy at Bologna, in experimenting on the muscular contractions produced by discharges from an electric machine, noticed that frogs' legs, hung on metal hooks in such a way that they rested in contact with a strip of another metal through which the hooks had been driven, were thrown into convulsive movements such as were produced by electric discharges. Following up the observation, he found that if strips of two unlike metals, such as zinc and copper, were taken and one put in contact with the main nerve of the frog's leg while the other was touched to the thigh muscles, spasmodic muscular contractions took place,

provided the other ends of the metal strips were in contact with each other.

606. Volta's Discovery. Volta, who was professor of physics in the University of Pavia, *believed that the source of the electrical effects observed by Galvani was to be found in the contact of the dissimilar metals.* But if there was any difference of potential produced in such a case it was far too small to be detected by the gold-leaf electroscope as ordinarily used. This difficulty was most ingeniously overcome by Volta's device of the *condensing electroscope.*

A gold-leaf electroscope was constructed having a flat brass plate instead of a knob, as shown in figure 331, on which rested a second brass plate of the same size having an insulating handle of glass by which it could be raised. Both plates were given a thin coating of shellac varnish by which they were insulated

FIG. 331. Volta's discovery

from each other and thus formed a condenser of large capacity, since the separating dielectric was thin.

The lower plate was then touched by a strip of copper soldered to the end of a zinc strip held in one hand while at the same time the upper plate was touched with the other hand. When the upper plate was raised after breaking these contacts, the gold leaves diverged with negative electricity, showing that the upper plate of the condenser had been charged positively and the lower negatively by the operation. The advantage of the condenser was that although the difference of potential between the plates was exceedingly small, a considerable charge was accumulated which was set free when the capacity of the condenser was decreased through separation of the plates.

When the two condenser plates were of brass and *directly connected* by the copper-zinc circuit, as in figure 332, no charge was obtained since the end metals were alike, being the two brass condenser plates; but if at any point in the circuit two dissimilar metals were connected by a dilute acid or salt solution, as shown in figure 333, the condenser plates were charged. In this case

the solution takes the place of the body of the experimenter in the original experiment.

It is now believed that the differences of potential obtained by Volta were mainly due not, as he supposed, to the contact of dis-
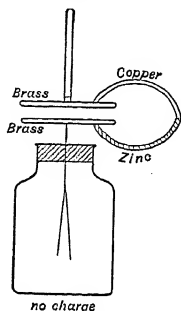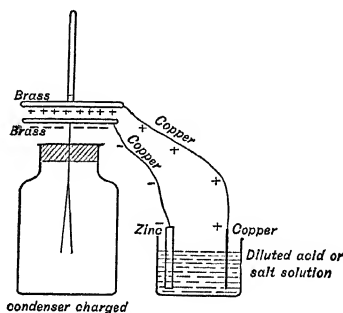
FIG. 332   FIG. 333

similar metals, but to the contacts between these metals and the hands of the experimenter or the acid or salt solution.

607. Voltaic Pile. In seeking to obtain a larger effect Volta found that when he took two cells in which strips of zinc and copper dipped into dilute acid, and joined them in *series*, as

FIG. 334. Charge by two cells

shown in figure 334, he obtained in the electroscope double the charge given by one cell. *The effect was found to depend only on the kind of metals and acid used and not at all on the size of the plates.*

The *Voltaic pile*, based on this discovery, consists of discs of copper, zinc, and cloth or paper saturated with acid or salt solution, piled one upon another, first a disc of copper, then acidulated cloth, then zinc, then again copper, cloth, and zinc, and so on. A pile having 50 such combinations will produce 50 times the difference of potential that can be obtained from a single element consisting of zinc-acid-copper.

What are known as *dry piles* are made by taking discs of gilt paper and so-called silver paper, placing them in pairs, the gilt face of one against the silver face of the other, and then

making a pile of such pairs, the same kind of paper being upper-most in each pair. In a moist climate the natural dampness of the paper enables it to play the part of the acidulated cloth layers in Volta's pile.

608. Voltaic Cell. A cell having a plate of zinc and a plate of copper dipping in dilute sulphuric acid is known as a simple Voltaic cell, and several cells combined constitute a Voltaic or Galvanic *battery*. Since Volta's day many improved kinds of battery cells have been devised, some of which will be considered later (§§ 638–646).

The two plates of a Voltaic cell are called the *electrodes*, and the terminals of the plates where the external wires are con-nected are called the *poles* of the cell. The copper terminal is at a higher potential than the zinc terminal and gives a positive charge, it is therefore called the *positive pole*, while the zinc terminal is the *negative pole*.

On the other hand, the copper plate is some-times spoken of as the *negative electrode* or *electro-negative element* in the cell and the zinc as the *positive electrode* or *electropositive element*, because electrons are transmitted through the acid of the cell from the copper to the zinc plate as though repelled by the copper and attracted by the zinc.
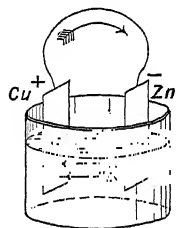
FIG. 335. Electric circuit

Electric Current in a Cell. Since the two poles of a Voltaic cell are at different potentials, an *electric current* is established when they are connected by a metallic wire just as when the two coats of a charged Leyden jar are connected. This current, however, flows steadily instead of lasting only for an instant. The metallic wire together with the plates and liquid between them form a conducting *circuit* in which *the positive direction of the current or that contrary to electron flow is from copper to zinc through the outside wire and from zinc back to copper inside the liquid of the cell.*

There are *three principal evidences* of the existence of the current:

1. Heat is developed in all parts of the circuit.

2. Every part of the circuit affects a magnetic needle brought near it.

3. Chemical action takes place at the surfaces of contact between the metal electrodes and the liquid. If the copper and zinc plates are in dilute sulphuric acid, bubbles of hydrogen gas appear at the surface of the copper plate, while the zinc plate is eaten away by the acid, and zinc sulphate is formed.

All these phenomena cease at once when the current is interrupted, either by breaking the metallic connection between the plates or by separating the acid around one plate from that around the other by a non-conducting partition.

609. Contact Potentials in a Voltaic Cell. When zinc is immersed in the acid there is what may be called a solution pressure, or tendency for the zinc to be dissolved and form zinc sulphate in solution, each atom of zinc carrying into the solution a positive charge.

As the positively charged atoms of zinc pass into solution, the plate, losing positive charge with each one, becomes negative, while the solution becomes positive, in consequence of which there is an electrostatic force tending to prevent the positively charged zinc atoms from going into solution. Therefore when a certain difference of potential between the zinc and acid solution is reached there will be equilibrium between the electrostatic force and the solution tendency, and the zinc will cease to be dissolved.

There is thus a definite difference of potential due to the contact of zinc and acid when there is equilibrium between them, and another due to the contact of copper and acid which is less than the former since the solution pressure of copper in the acid is less than that of zinc.

610. Electromotive Force. The diagram of figure 336 represents the relative potentials of the elements in a Voltaic cell. The liquid potential is uncertain because of difficulties in assigning a definite potential to the acid.

If the copper pole of the cell is connected to the earth, it comes to the earth potential or zero, and the zinc pole as tested by a quadrant electrometer is found to have a negative potential. On the other hand if the zinc pole is connected to the earth it will be at zero potential while the copper pole will be found to be positive; but *the difference of potential* between them will be the same in each case.

Every cell can produce a certain maximum difference in potential between its two electrodes, and when this is reached there is equilibrium and the chemical action stops.

The maximum difference of potential which a cell can produce is called its electromotive force; it is measured by the difference of potential between the electrodes when there is no current and the chemical action has ceased.

The electromotive force of a cell depends only on the chemical relations of the constituents of the cell and is therefore the same whether the plates are large or small.

Fig. 336

A small cell formed by dipping the tips of a zinc and of a copper wire into a single drop of acid will cause as great a deflection of a quadrant electrometer as a cell of the same kind with plates a foot square.

A convenient abbreviation for electromotive force is E.M.F., or in equations the symbol E is commonly used.

611. Hydraulic Analogy to Voltaic Cell. The following analogy given by Lodge is instructive. Two tall open vessels con-
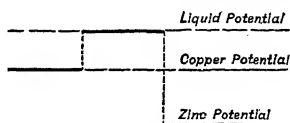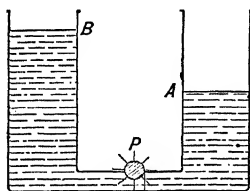
Fig. 337
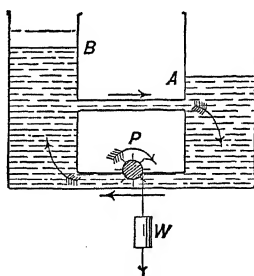
Fig. 338

taining water are connected by a pipe in which is a pump *P* driven by a weight *W* (Fig. 337). The water will flow from one vessel to the other until the back pressure on the pump due to the higher level of *B* just balances the force of the weight. The difference in level will be the same whether the vessels are large or small. The difference of level represents the difference of potential between the zinc and copper which is independent of the size of the cell, the pump with its driving weight is the

electromotive force of the cell, which through chemical action can produce a certain definite difference of potential and no more.

Figure 338 represents the state of things when the zinc and copper plates are connected by a wire, represented by the tube shown. The difference of pressure causes a flow through the tube from B to A, at the same time the level sinks in B and rises in A so that the difference in pressure on the two sides diminishes and is no longer able to balance the pressure of the pump, which therefore begins to act, forcing water from A to B; at the same time the weight W descends, supplying energy for the circulation, which will be maintained so long as the weight can move downward.

Here it is seen that the electromotive force, represented by the power of the pump to produce pressure, is the same as before, but the difference of potential between the plates, shown by the difference between the levels of A and B is less than before. The work done by the pump in circulating the water is obtained from the weight, which loses potential energy as it descends. So in the Voltaic cell, *the energy expended by the electric current is supplied by the chemical changes which take place at the electrodes*.

612. Magnetic Effect of Current. In 1819 Oersted discovered that when a wire connecting the poles of a Voltaic cell was held *over* a balanced magnetic needle and parallel to it, the needle was deflected, the north pole of the needle moving toward the west when the current was from south to north, as in the diagram, while if the current was reversed the north pole of the needle moved toward the east. The effect was reversed when the wire was placed *under* the needle.

This discovery aroused the greatest interest, as *it was the first evidence of a connection between magnetism and electricity*.

613. Electric Circuit. It was also found that the action was the same whatever part of the wire connecting the plates was brought near the needle, the deflection produced by the current in the middle of the wire being just as great as that near its ends.

By this experiment also the *direction of the current in the electrolyte may be shown to be opposite to that in the wire;* for if two vessels are used connected by a short tube containing the acid, and if a zinc plate is placed in one vessel and copper in the

other, as shown in figure 340, a magnetic needle will be deflected toward the west when placed under the wire connecting the plates, but toward the east when under the tube. The experiment shows that the current in the electrolyte is just as strong as that in the wire, but in the opposite direction.

From experiments such as the above it is inferred that *steady electric currents always flow in closed circuits and are equally strong at every point*, and if the circuit is interrupted at any point, whether in the electrolyte or the wire, the magnetic action and all other current effects cease everywhere at almost the same instant. It is very much as when an incompressible liquid circulates in a closed tube, just as much liquid must pass any one section of the tube as any other during the same time.
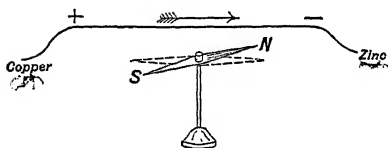
FIG. 339. Current and magnetic needle   FIG. 340. Current in electrolyte
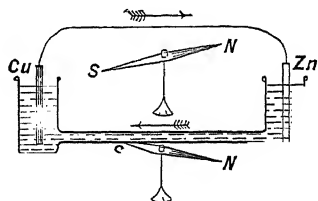
614. Galvanometers. When a wire is bent into a vertical circle having its plane parallel to the direction of a magnetic needle balanced at its center, if a current is established in the wire all parts of it act together to deflect the needle, turning the north pole to one side or the other, depending on the direction of the current. An instrument which measures electric currents by the deflection of a magnetic needle is known as a *galvanometer*.

615. Galvanometers Measure Current. *Faraday showed that a galvanometer measures the quantity of charge transmitted per second, or what is called the current strength.* For he found that when a Leyden jar was discharged through a sensitive galvanometer there was an instantaneous swing of the needle to one side, the amount of which depended only on the *quantity* of the charge; that is, the swing produced by forty turns of his electrical machine was the same whether the charge was held in a small jar at high potential or in a large Leyden battery at low

potential, and whether the wet string through which the discharge was sent was long or short.

It was also established by Faraday that when *a constant current* flowed through a galvanometer producing *a steady deflection* of the needle, *the magnetic force on the needle due to the current was proportional to the quantity of charge transmitted per second.*
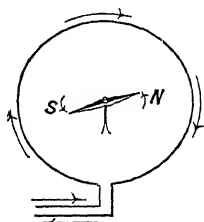
FIG. 341. Galvanometer

616. Unit Current. Instead of measuring electric currents by the quantity of charge in electrostatic units transmitted per second, it is found better to adopt a new system of units based on the magnetic effect of a current and using magnetic units as already defined. This system is known as the *C. G. S. electromagnetic system,* since it also is based on the centimeter, gram and second.

In this system *a unit current is one which, flowing in a circular coil of one centimeter radius, will act on a unit magnetic pole at its center with a force of one dyne for every centimeter of wire in the coil.*

The Ampère or Practical Unit of Current. The unit of current in the practical system is called the *Ampère* in honor of the French physicist who first investigated the laws of the magnetic effects of currents. It is defined as *one-tenth of the C. G. S. electromagnetic unit current,* being chosen smaller than the latter unit for reasons of convenience.

The quantity of charge transmitted by one ampère in one second is called a *coulomb.* One coulomb is equal to 3,000,000,000 electrostatic unit charges as defined in § 540.

617. Unit of Electromotive Force. In our studies of electrostatics it was shown (§ 571) that the difference between the potentials of two conductors might be measured by the work required to transfer unit charge from one conductor to the other. Just so in the *absolute electromagnetic system of units two points in a conductor are said to have unit difference of potential when one erg of work is required to transfer the C. G. S. electromagnetic unit quantity of electricity from one point to the other.*

Unit quantity of electricity in the C. G. S. electromagnetic system is of course the charge transmitted per second by unit current in that system.

The unit of potential in the electromagnetic system is found to be so small compared with the electromotive forces of ordinary battery cells that it was decided to adopt for ordinary use a unit one-hundred million times as great, called the *volt* in honor of Volta.

The volt is the unit of electromotive force in the practical system and is 10^8 times as great as the C. G. S. electromagnetic unit of potential. It is much smaller than the electrostatic unit of potential defined in § 571, the latter being almost exactly equal to 300 volts.

The electromotive force of the Voltaic cell is nearly 1 volt.

618. Resistance. Let a circuit be made up of two battery cells A and B joined in series with some other conductors and a galvanometer, the two cells being so connected that their electromotive forces act in the same direction. After observing the current strength as shown by the galvanometer, let the circuit be rearranged, taking the same components in any other order whatever. If the two electromotive forces still act together the current will be found the same as before, showing that the current strength is not affected by the particular order of the parts in an electric circuit. But if one cell is turned around so that its electromotive force opposes that of the other cell, then the effective electromotive force in the circuit will be the *difference* between the electromotive forces of the two cells instead of their *sum* as in the former case, and the current in this case will be smaller than before, just in proportion as the electromotive force is smaller.
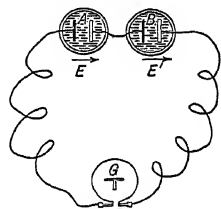
FIG. 342

That is, the current strength is proportional to the effective electromotive force; or, in other words, *the ratio of the electromotive force to the current strength in a given circuit is a constant,* which depends only on the make-up and physical condition (temperature, stress, etc.) of the circuit. *This constant is called the resistance of the circuit and is not affected by the order in which the various conductors, cells, etc., are connected, nor by the direction in which the current flows through them.*

This relation, established by the German physicist, G. S. Ohm, is known by his name and may be stated as follows:

Ohm's Law: *The ratio of electromotive force to current in a given circuit is a constant which may be called the resistance of the circuit.*

Or, in symbols, $\dfrac{E}{I} = R = a\ constant$

where E represents the electromotive force, I the current, and R the resistance of the circuit.

It is also established by experiment that *each battery cell and piece of wire or other conductor has a definite resistance which belongs to it individually and depends only on its temperature and state of stress* (provided that the same two points on the conductor are always used in making connection with the rest of the circuit); and when the several parts of a circuit are joined together one after another, in *series* as it is called, the resistance of the whole is the sum of the resistances of the several parts.

619. Unit of Resistance. In honor of the discoverer of this law the unit of resistance in the practical system is called the *ohm;* it is *the resistance of a circuit in which an electromotive force of one volt will produce a current of one ampère.*

Ohm's law may then be expressed in units of the practical system, thus:

$$Current\ in\ ampères = \frac{Electromotive\ force\ in\ volts}{Resistance\ in\ ohms}.$$

The electrical resistance of a conductor is analogous to the frictional resistance which a pipe offers to the flow of liquid through it. In both cases work done against the resistance appears as heat, and in neither case does the resistance have any tendency to produce a back current.

620. Exception to Ohm's Law. In gaseous conductors the ratio of the electromotive force to the current is not constant as in other conductors, but depends on the strength of the current.

CHEMICAL EFFECTS OF CURRENT

621. Decomposition of Water. When a current of electricity is passed through dilute sulphuric acid (1 part acid to 10 of water), using platinum electrodes immersed in the acid, gas is given off at each electrode. The gases may be separately collected in tubes filled with the dilute acid and inverted over the electrodes as

shown in figure 343. The gas liberated at the positive electrode is found to be oxygen while that at the negative electrode is hydrogen, and the volume of hydrogen is just twice the volume of the oxygen. These volumes are exactly in the ratio in which the gases combine to form water, and on this account it was at first supposed that the current directly decomposed water.

The decomposition of water in this way by the electric current was first accomplished in 1800 by Carlisle and Nicholson.

622. Discovery of Potassium and Sodium. Sir Humphrey Davy, in 1807 by the use of a powerful battery of 250 cells, decomposed caustic potash, obtaining metallic potassium at the negative electrode. A fragment of caustic potash slightly moistened was laid on a platinum plate which was connected to the positive pole of the battery; on touching the potash with a platinum wire connected with the negative pole, minute globules appeared at the negative electrode which rapidly oxidized in air or took fire; these he recognized as a new metal which he named potassium. In a similar manner metallic sodium was obtained from caustic soda.
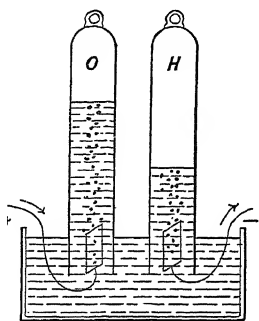
Fig. 343. Electrolysis of water

623. Faraday's Researches. About the year 1833 Faraday began the systematic investigation of the chemical effects of the electric current.

Substances which are decomposed by the passage of the electric current he called *electrolytes;* the electrode connected with the positive pole of the battery or that through which (according to ordinary convention) current enters the electrolyte was named the *anode* (Greek, *inward path*), while the electrode through which the current leaves the electrolyte was named the *kathode* (Greek, *outward path*). The two constituents into which a molecule of the electrolyte is broken up were called *ions* (Greek, *wanderers*), that which is set free at the kathode being the *kation*, while that which appears at the anode was named the *anion.*

624. Faraday's Laws. The following are two important results of Faraday's investigations:

1. The quantity of a given substance electrolyzed in a cell is

proportional to the amount of charge or quantity of electricity which passes.

2. If several electrolytic cells containing different substances are connected in series in the same circuit, the quantities of the ions set free at the electrodes are proportional to their chemical combining equivalents.

625. Electrochemical Equivalents. *The electrochemical equivalent of a substance is the quantity that is set free per second by a current of one ampère* or by the passage of one *coulomb* of electricity. The following table gives the electrochemical equivalents of some well-known substances. It will be noticed that they are proportional to the combining equivalents.

ELECTROCHEMICAL EQUIVALENTS

SUBSTANCE	ATOMIC WEIGHT	VALENCE	COMBINING EQUIVALENT	ELECTROCHEMICAL EQUIVALENT (GMS. PER COULOMB)
Kations				
Hydrogen.....	1	1	1	0.000010357
Copper........	63.18	2	31.59	0.00032840
Silver.........	107.7	1	107.7	0.00111800
Anions				
Oxygen.......	16	2	8	0.00008283
Chlorine.......	35.37	1	35.37	0.0003671

96,550 coulombs are transmitted when the number of grams liberated equals the combining equivalent of the substance.

626. Primary and Secondary Actions. It is important to distinguish between the direct or primary effect of the current in electrolysis and the secondary chemical reactions that take place when the ions are set free. In illustration of this difference take the electrolytic apparatus containing dilute sulphuric acid, as described in § 621, and connect it in series with a precisely similar apparatus containing a solution of sodium sulphate in water, colored by an infusion of purple cabbage. On sending a current through, both hydrogen and oxygen gases are set free in one cell exactly as in the other, and at the same time the coloring matter in the sodium sulphate solution turns red around the positive electrode, or *anode*, and green around the negative electrode, or *kathode*, showing that the originally neutral salt has become acid

at the anode and alkaline at the kathode. Analysis shows that sodium hydroxide (NaOH) has appeared at the one electrode and sulphuric acid (H_2SO_4) at the other.

It might seem at first that more decomposition was effected by the current in one cell than in the other, in violation of Faraday's law, for equal amounts of gas are set free in both cells, and in addition the sodium sulphate in the second cell is decomposed, while the sulphuric acid in the first cell remains unchanged.
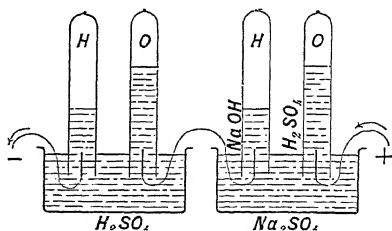
FIG. 344. Electrolytic cells in series

But it is believed that the *primary* effect of the current is to separate precisely equivalent quantities of H_2SO_4 and Na_2SO_4 in accordance with Faraday's law, the other changes being *secondary* chemical actions.

Thus in the sulphuric acid cell the primary action of the current is to separate the H_2 and SO_4 ions; oxygen (O) is set free at the anode as the result of a secondary reaction in which the SO_4 ion displaces the oxygen from a water molecule (H_2O) and forms sulphuric acid (H_2SO_4), which remains in solution.

In sodium sulphate the primary action of the current is to separate the Na_2 and SO_4 ions. Then secondary reactions take place at both electrodes, the SO_4 ions effect the liberation of oxygen at the anode exactly as in the other cell, while the positively charged sodium (Na_2) ions pass to the kathode, where each combines with two molecules of water $2(H_2O)$, forming sodium hydroxide $2(NaOH)$, which remains in solution, and setting free hydrogen (H_2), which gives up its positive charge and escapes at the kathode.

These secondary reactions may be expressed symbolically thus:

$$SO_4 + H_2O = H_2SO_4 + O$$
$$Na_2 + 2(H_2O) = 2(NaOH) + H_2.$$
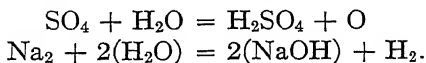
627. Theory of Electrolysis. The earlier explanations of electrolysis supposed the decomposition of the electrolyte to be effected by the electric current, but it is now believed that a

large per cent of the electrolyte is *ionized*, or broken into positively and negatively charged ions, as a result of going into solution, and that the electric force in the electrolyte is simply *directive*, causing the positively charged ions to move with the current and the negatively charged ions to move in the opposite direction through the solution until they reached the electrodes where their charges are given up and the molecules are set free in the neutral state. The current is supposed to be made up of the charges which are thus carried convectively by the moving ions.

Hittorf showed that different kinds of ions moved through the electrolyte with widely different velocities, and measured the relative velocities of anions and kations in aqueous solutions of many different salts and acids. Kohlrausch, by measurement of the electric charges transmitted per second through these solutions, and assuming that the electric current is transmitted through the electrolyte wholly by the charges carried by the moving ions, has determined the actual velocities with which different kinds of ions move in aqueous solutions, and finds them proportional to the electric force, that is, to the fall of potential per centimeter in the solution.

Some values found by Kohlrausch are given in the following table.

Ionic velocities for a potential gradient of one volt per centimeter

Kations			Anions		
Na	45.	$\times 10^{-5}$ cm./sec.	Cl	69.	$\times 10^{-5}$ cm./sec.
H	320.	" "	NO^3	64.	" "
Ag	57.	" "	OH	182.	" "

628. Ionic Charges. Faraday's laws show that every univalent ion carries a certain charge $\pm e$ which is either positive or negative, depending on whether the ion is an anion or a kation; while bivalent and trivalent ions carry charges $\pm 2e$ and $\pm 3e$, respectively.

It was suggested by Helmholtz that the charge e may be the *atom of electricity* from which all other charges are made up and of which they are therefore multiples. This is borne out by the experiments of Millikan as we have already seen (§ 560).

Like all other electric charges, these ionic charges are pro-

duced by electrons. Positive ions are those which have lost electrons.

In the electrolysis of 1 gram of hydrogen 96550 coulombs of electricity are transmitted; and assuming that each atom of hydrogen carries the charge e equal in magnitude to that of the electron, we find that in 1 gram of hydrogen there are 6.06×10^{23} atoms.

629. Polarization. At the electrodes where the ions are set free or enter into new combinations there are generally electromotive forces, because at those points electric energy has to be spent to effect chemical changes. The resultant of these electromotive forces is called the polarization of the cell.

In case of the electrolysis of copper sulphate between copper electrodes the chemical change which takes place at the kathode is opposite to that at the anode: Cu and SO_4 are separated at the one and united at the other. Therefore, in such a cell there is on the whole no electromotive force of polarization.

But if dilute sulphuric acid is electrolyzed between platinum electrodes there is an electromotive force developed *against* the current, or a back electromotive force of about 1.7 volts, and unless the battery employed has an electromotive force greater than this the current cannot be maintained.

While the electrodes are thus polarized the cell is in reality a *battery cell*, and if it is disconnected from the main circuit and its electrodes joined by a conducting wire, a current is obtained opposite to that which caused the polarization. This current flows until the gaseous layers on the electrodes disappear. The cell is thus really a storage battery cell of very small capacity.

All storage battery cells or accumulators depend on the electromotive force of polarization.

When dilute sulphuric acid is electrolyzed with a zinc anode and copper kathode, as in the simple Voltaic cell, more chemical energy is given out at the anode where zinc sulphate is formed than is absorbed at the kathode where hydrogen is liberated from the solution, and consequently, on the whole, energy is given out by the chemical changes instead of being required to bring them about, hence the electromotive force of polarization is *with the current* instead of against it, and the combination is called a battery cell.

630. Measurement of Current. Currents of electricity are conveniently measured by their electrolytic effect. In the instrument shown in figure 345,

FIG. 345. Voltameter or Coulomb-meter

known as a *voltameter*, dilute sulphuric acid is electrolyzed between platinum electrodes, and the escaping gases are caught mingled together in the graduated tube above the electrodes. From the temperature, volume, and pressure of the collected gas its weight can be determined, and if the time during which the current was flowing is known the current can be calculated, since 1 ampère will set free in 1 minute 0.00559 gms. or about 12.2 c.c. of the mixed gases.

631. Copper Voltameter. A more accurate instrument for the measurement of current is the copper voltameter, in which copper sulphate is electrolyzed between copper electrodes. From the gain in weight of the kathode while the current is flowing the amount of copper deposited per second is determined, and so the current is found from the electrochemical equivalent of copper. The form shown in figure 346 is convenient. The alternate plates are connected into one set and form the anode, while the intermediate plates form the kathode, so that each kathode plate is between two anode plates. The number of plates used depends on the strength of the current to be measured. To secure the best results something like 40 sq. cms. surface per ampère is required in the kathode.

632. Silver Voltameter. For standard determinations it is found that the most reliable results are obtained from a form of *silver voltameter* in which the liquid is a standard solution of nitrate of silver contained in a platinum cup which also serves as the kathode, while the anode is a rod of pure silver which dips into the liquid. The anode must be surrounded by a covering to prevent any particles of silver that may become loosened from the anode from falling into the platinum cup.

633. Electroplating. By means of the electric current metallic objects may be plated with gold or silver or other metals. Figure 347 shows a form of electroplating bath. The objects to be plated are hung on metal rods which are all connected with the negative pole of the battery or dynamo

FIG. 346. Copper voltameter

which supplies the current. If silver is to be deposited, plates of silver connected with the positive pole are hung in the bath between the objects which are being plated, so that while silver is being deposited the anode plates lose an equal amount and the strength of the solution is maintained

constant. The thickness of the deposit will generally be greater on projecting parts of the object plated and on parts that are nearer to the anode plate.

PROBLEMS

1. If in a copper voltameter the kathode plates gain 1.50 grms. of copper in ten minutes, find the average current strength in ampères.

2. How many grams of hydrogen and of oxygen are set free when 1 grm. of water is decomposed by electrolysis; and how many cubic centimeters of each of these gases will there be at 0° C. and 76 cms. pressure?

3. How many coulombs of electricity will be required to effect the decomposition in the previous problem, and how long a time must a current of 0.5 ampère flow to accomplish it?
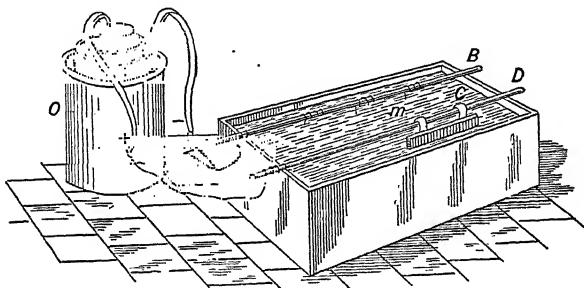
FIG. 347. Electroplating bath

BATTERY CELLS

634. Battery Cells. A battery cell is a combination in which electromotive force is produced by chemical action. The simple cell of Volta is the earliest type, but it has important practical defects.

An ideal cell will have:

1. Small resistance.
2. Large electromotive force.
3. A constant electromotive force whatever the current.
4. No local action or wasteful chemical action.

635. Resistance of Battery Cells. When the electrode plates are large and close together the resistance of the cell is small; while if the plates are very small the resistance of the cell may be so great that even when the poles are *short-circuited* or con-

nected by a short copper wire offering very little resistance, the current will be extremely small.

Cells from which large currents are to be obtained must, therefore, have large plates separated by a comparatively thin layer of electrolyte.

636. Local Action. If commercial zinc is used in a Voltaic cell hydrogen gas will be given off at the surface of the plate as soon as it is placed in the acid and before it is connected with the copper plate. This is accompanied by a corresponding wearing away of the zinc and formation of zinc sulphate, which goes into solution. This wasting of the zinc is called *local action* and is due to impurities. Suppose that a particle of iron or carbon imbedded in the surface of the zinc is in contact both with the zinc and acid; it forms a minute Voltaic cell, in which the current flows from the iron or carbon to the zinc and through the acid from zinc to iron again, as indicated in the figure, and zinc is eaten away near the impurity and hydrogen set free at its surface.
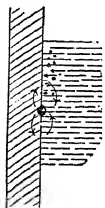
Fig. 348. Local action

To prevent local action the zinc surface is freshly amalgamated with mercury, which dissolves the zinc, covers up the impurities, and presents a homogeneous surface to the acid.

637. Polarization. When the poles of a simple Voltaic cell are connected by a wire, the current does not remain constant but rapidly decreases in strength.

This weakening of the current is due to *polarization*. The hydrogen set free at the copper electrode forms a sort of gaseous layer over the plate which interferes with the action of the cell in two ways. In the first place, *the resistance of the cell is increased*, for the flow of electricity is interfered with by the bubbles of gas. In the second place, *the electromotive force of the cell is diminished*, for the hydrogen layer is much more like zinc in its relation to the acid than is the copper which it covers.

This difficulty is most effectively met by the use of two electrolytes.

638. Primary and Secondary Battery Cells. Cells such as the Voltaic cell in which the current is obtained from the chemicals of which the cell was originally constructed are known as *primary cells*, while cells in which the chemical state necessary for the production of a current is produced by sending through the cell a current from some outside source for a cer-

tain length of time, are known as *secondary batteries, storage cells,* or *accumulators.*

A few of the cells most commonly used in practice will now be considered.

Primary Battery Cells

639. The Daniell Cell. One of the first and most useful two fluid cells was devised by Daniell in 1836. It consists of a copper electrode immersed in a solution of copper sulphate and an electrode of amalgamated zinc immersed in dilute sulphuric acid, the two being separated by a partition of porous earthenware. In figure 349 the copper electrode with its solution is represented as contained in a cup of porous earthenware surrounded by the zinc and dilute acid.
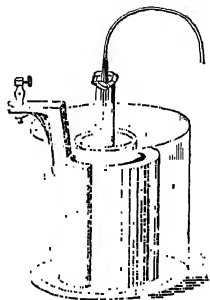
FIG. 349. Daniell's cell

When the circuit is closed, the positively charged zinc atoms pass into solution forming zinc sulphate with the negative SO_4 ions, while the positively charged hydrogen ions (H_2) in the acid move toward the copper plate, passing through the porous cup by diffusion and forming sulphuric acid (H_2SO_4) with the negative SO_4 ions from the copper sulphate, and displacing the positive copper ions (Cu) which give up their charges and are deposited on the copper plate.

In the dilute acid $\qquad$ $Zn + H_2SO_4 = ZnSO_4 + H_2.$

In the copper sulphate $\qquad$ $H_2 + CuSO_4 = H_2SO_4 + Cu.$

Thus zinc is dissolved and zinc sulphate formed, copper sulphate is used up and copper deposited on the copper electrode. There is no hydrogen layer formed on the copper and consequently no polarization. The electromotive force of this cell is about 1.08 volts.

640. Gravity Cell. A form of Daniell cell which has been extensively used in telegraphy and is still much used where a *small constant current* of electricity is required is the *gravity cell,* so called because the liquids are kept separate by gravity alone, the denser copper sulphate solution resting at the

bottom of the cell, while the lighter acid or zinc sulphate solution floats above it.

If the gravity cell stands without being used the copper sulphate diffuses gradually up into the acid above and copper is deposited on the zinc, causing extensive local action. A small current, sufficient to balance the diffusion, should always be kept flowing while the cell is set up.

641. Leclanché Cell. This very useful form of cell has a zinc and a carbon electrode. The carbon is packed in a porous cup with a mixture of fragments of carbon and black oxide of manganese; the zinc electrode is in a strong solution of ammonium chloride (sal ammoniac) which surrounds the porous cup.

The hydrogen which would polarize the carbon electrode combines with oxygen from the manganese dioxide and forms water. But as the depolarizing agent is in the solid form its action is slow, and the cell polarizes temporarily. It is extensively used, however, for *open-circuit* work, such as for bells, annunciators, and clocks, where a steady current is not required. It is entirely free from injurious or disagreeable fumes, there is but little local action, and no trouble from diffusion, so that the cells may stand set up for a year or two without attention, and ready for use at any instant. Its electromotive force is about 1.40 volts.

FIG. 350. Gravity or crow-foot cell

The chemical changes in this cell are:

$$2NH_4 + MnO_2 = 2NH_3 + H_2O + MnO$$

$$2Cl + Zn = ZnCl_2.$$
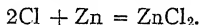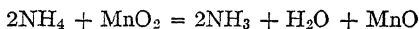
642. Dry Cells. The so-called dry cells are ordinarily a form of Leclanché cell. The outer cylindrical cup forms the zinc electrode which is lined with thick absorbent paper and packed with the pulverized manganese dioxide and carbon mixture surrounding the central carbon rod. The whole is saturated with ammonium chloride solution and sealed with pitch to keep it from drying out.

FIG. 351. Leclanché cell

Secondary Cells

643. Grove's Gas Battery. The English physicist Grove showed that when in the decomposition of water long electrodes

were used, extending to the tops of the tubes in which the gases were collected, as in figure 352, on changing the switches ss' to the dotted positions, thus disconnecting the battery B and simply joining the two electrodes together through a galvanometer G, a current was obtained which was in the opposite direction to the decomposing current. At the same time *a gradual recombination of the hydrogen and oxygen took place until these gases had entirely disappeared.*
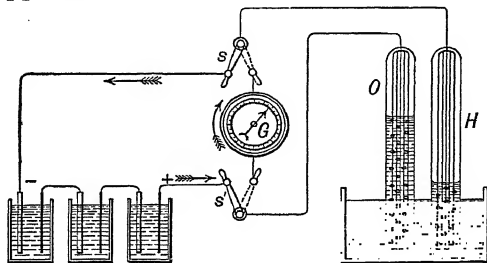
Fig. 352

Long electrodes were necessary since each electrode must pass through the surface where the gas and electrolyte meet.

644. Planté Cell. If a current of electricity is sent through a cell consisting of two plates of sheet lead in dilute sulphuric acid it becomes polarized, one plate becoming oxidized while hydrogen is set free at the surface of the other, reducing any oxide that may be there.

In the year 1860 Planté, a French physicist, found that secondary cells of large capacity could be made in this way. His method was to send a current through the cell in one direction until one plate was well oxidized, after which the cell was discharged and then charged by a current in the opposite direction, thus oxidizing the other plate and reducing the oxide on the first to metallic lead in a spongy form. The cell was then again discharged and charged with a current in the same direction as at first, and so by alternately charging and discharging, first making one plate positive and then the other, a deep layer of active material was formed on each plate. The plates were then said to be *formed*.

645. Storage Cells — Accumulators or Secondary Batteries. Secondary battery cells, or storage cells as they are frequently called, have become extensively used in electric motor vehicles, in electric power plants, and in telegraphy. The plates are usually heavy lead grids full of holes or grooves containing the active material, which is either packed in them mechanically

or formed in them by some such process as that used by Planté. The positive plates contain a high oxide of lead, PbO_2, while the active parts of the negative plate are of spongy lead. A cell is formed of a number of such plates, alternately negative and positive, as shown in the figure, set in a suitable vessel containing dilute sulphuric acid. The negative plates are connected together and form one set and the positive plates form another, there being one more negative than positive plate, so that each positive plate is between two negative ones. This is to prevent buckling or bending of the positive plate, for the formation of oxide in charging is accompanied by an increase in volume or swelling of the plate, which would warp it badly if it took place only on one side.

The nearness of the plates to each other and the large surface obtained by using a number of plates cause the resistance of the cell to be very small. The greater the number and size of the plates in a cell the larger the current that can be sent through it without injury to the cell. About 1 ampère per 12 sq. in. of opposed surface is usually a safe rate of discharge.

FIG. 353.　Storage cells

The commercial importance of such storage cells is due in part to their *extremely small resistance* and to the fact that they are renewed not by means of costly chemicals, but by a current obtained from a dynamo machine driven by an engine or by water power. They can be used, therefore, to store the superfluous energy of a power plant at times when but little power is used and give it back again in times of need.

The electromotive force of this type of storage cell is about 2.10 volts. The chemical changes in such a cell are as follows:

Discharging

Positive plate $\qquad PbO_2 + H_2 + H_2SO_4 \qquad PbSO_4 + 2H_2O_4$

Negative plate $\qquad\qquad\qquad Pb + SO_4 \qquad PbSO_4.$

Lead sulphate is thus formed at each plate.

Charging

Positive plate $\qquad PbSO_4 + SO_4 + 2H_2O = PbO_2 + 2H_2SO_4$

Negative plate $\qquad\qquad PbSO_4 + H_2 = Pb + H_2SO_4.$

646. Edison Storage Cell. A form of storage cell has been devised by Edison in which the active materials of the electrodes are the oxides of nickel and iron, respectively, the electrolyte being a solution of caustic potash in water. A battery of these cells weighs about one-half as much as the equivalent lead cells. The cells are very durable and are not so easily

injured as lead cells by overcharging or leaving uncharged. Its disadvantages are a lower and less constant voltage and a much greater temperature coefficient than exist with the lead storage cell.

Modes of Connecting Cells

647. Battery Cells in Series. When battery cells are connected as shown in the figure, the positive pole of one being joined to the negative pole of the next, they are said to be joined in *series*, and *the electromotive force of the combination is the sum of the electromotive forces of the several cells.* As the whole current must pass through each cell *the resistance of cells joined in series is the sum of the resistances of the separate cells.*
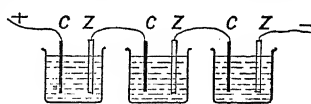
FIG. 354

Suppose that three cells, each with electromotive force e and resistance r, are connected in this way with an external resistance R. The total electromotive force is $3e$ and the resistance is $3r + R$ so that by Ohm's law,

$$I = \frac{3e}{3r + R} \qquad \text{(for 3 cells in series)}$$

where I is the current.

This arrangement is advantageous when the external resistance R is large and a large electromotive force is required.

Battery cells in series may be likened to a series of pumps, the first of which lifts water to a certain level where the second takes it and lifts it to the next higher level and then the third raises it again to a still higher level, etc.

648. Battery Cells in Parallel. If cells are joined together as shown in figure 355, all the copper poles being connected together for the positive pole and all the zincs for the negative, they are said to be joined *in parallel.*
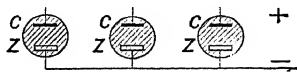
FIG. 355

Such a combination has precisely the advantage that a large cell has over a small one. *Its electromotive force is the same as that of one cell, while its resistance is less* — a combination of four similar cells joined in this way having only one-fourth the resistance of a single cell.

Only similar cells should be joined in parallel, otherwise a cell of smaller electromotive force may have a reverse current sent through it by a stronger cell.

This mode of arrangement is useful when the external resistance in the circuit is much smaller than that of a single battery cell or where the current to be obtained is more than can advantageously be transmitted through a single cell.

Cells in parallel may be likened to a set of pumps which are all lifting water from the same lower canal to another at a higher level.

649. Combined Series and Parallel Arrangement of Cells. Several similar series of cells may be combined in parallel as

FIG. 356

shown in figure 356, where two series of three cells each are connected in parallel. It will be observed that the resistance of each of the rows is $3r$ and the electromotive force of each row is $3e$. The resistance of the two rows in parallel is then $\dfrac{3r}{2}$ or one-

half that of the row, while the electromotive force of the combination is the same as that of a single row. The expression for current is then, for this particular arrangement,
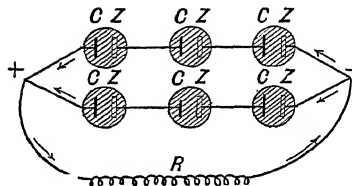
$$I = \frac{3e}{\dfrac{3r}{2} + R}$$

A square arrangement, where the number of cells in each row is the same as the number of rows, will have the same resistance as a single cell.

If a given number of cells are to be combined so as to give the maximum current through a given outside resistance, use that arrangement which will make the battery resistance most nearly equal to the external resistance.

Commercial types of storage battery cells have such small resistances that except where very large currents are required they are connected in series. If large currents are to be obtained, several series of cells are arranged in parallel so that the current

through each series will only be such as the cells are adapted to transmit without injury, and in each series as many cells are used as are necessary to give the required electromotive force.

PROBLEMS

1. If two gravity cells and one Leclanché cell are joined in series with a coil of wire having a resistance of 5 ohms, what current is obtained, when the gravity cells have a resistance of 2 ohms each, while the Leclanché cell has resistance 0.4 ohm?

2. Find the current when the zinc pole of a gravity cell of 2 ohm resistance is connected to the zinc pole of a Leclanché cell of 0.5 ohm resistance while their other poles are connected by a wire of 1 ohm resistance.

3. Reverse the gravity cell in problem 2 so that its copper pole is connected to the zinc pole of the Leclanché cell, the other poles being connected by the 1-ohm wire, and find the current as before.

4. Make a diagram of two gravity cells of 2 ohms resistance each, connected in parallel to a coil of wire having 1 ohm resistance, and show what is the current in the wire and what is the current through each cell.

5. What is the electromotive force and internal resistance of a combination of 12 gravity cells, consisting of three series of four cells each, the three being connected in parallel? Take resistance of each cell 2 ohms and its E.M.F. 1 volt.

6. When the terminals of the battery described in problem 5 are connected by a wire having a resistance of 3 ohms, find the current in the wire and also in each cell.

7. If a single storage cell has an electromotive force of 2 volts and a maximum permissible discharge rate of 10 ampères, how many such cells will be required and how arranged to give a current of 30 ampères and have an electromotive force of 50 volts?

8. If the cells in the last problem each have a resistance of 0.01 ohm, find what is the smallest resistance that can be permitted in the outside circuit.

9. How many gravity cells having a resistance of 2 ohms and E.M.F. 1 volt each, will be required to light a 50-volt incandescent 16-candle-power lamp which has a resistance of 50 ohms and requires 1 ampère of current, and what arrangement will require the smallest number of cells?

NOTE: The smallest number will be required when the battery resistance is equal to the external resistance.

FALL OF POTENTIAL AND RESISTANCE

650. Fall of Potential along a Circuit. In every electric circuit there is a gradual decrease of potential along the external circuit from the positive to the negative pole of the battery.

Faraday showed that *in all parts of a simple undivided circuit the current is of the same strength.* Therefore, except at points where electromotive forces are introduced, the potential must everywhere gradually change from point to point, for there is no current between points at the same potential. In figure 357 is shown a circuit with a corresponding diagram of potentials, the potential at each point of the circuit being represented on the diagram by the vertical distance from the base line to the upper curve. Starting at C, it will be observed that the potentials steadily diminish till the zinc pole is reached at Z, where the lowest value is found. But passing from the zinc
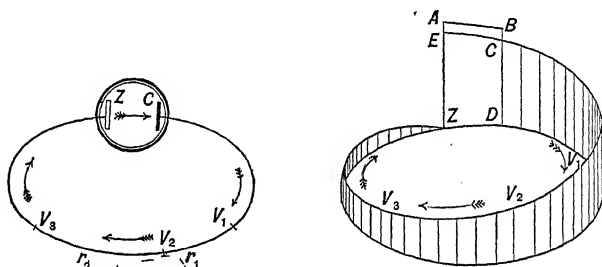
FIG. 357. Circuit and diagram of potentials

into the acid the potential is raised to its maximum at A by the chemical action; here there is an electromotive force and consequently a sudden change in potential. In the acid the potential again steadily falls as we pass from zinc to copper, until at the surface of the copper there is another electromotive force and a consequent sudden drop in potential bringing us again to the starting point at C.

If in the diagram AE is laid off equal to BC, then EZ is the difference of potential which represents the total electromotive force of the cell, for it is the difference between the two oppositely directed electromotive forces AZ and BC, and the fall of potential from A to B is equal to that from E to C. It is therefore clear that *the total fall of potential around the circuit, including that which takes place within the acid as well as that in the outside circuit, is equal to the total electromotive force of the cell.*

It will be noticed that CD, *the difference of potential between the two poles, is equal to the fall of potential in the external circuit,*

and *is therefore less than the total electromotive force of the cell when-ever there is any current flowing,* for there is then a fall of potential within the cell itself from A to B.

651. The Potentiometer. A very useful way of comparing potentials is by means of the potentiometer, a simple form of which is shown in figure 358. A wire of high resistance AB has the battery C connected to its terminals so that a steady current flows in it. If the wire is uniform, there is a uniform fall of potential along it from the end connected to the positive side of the battery to that connected to the negative side just as shown in figure 357. If the negative pole of a cell D, whose potential must
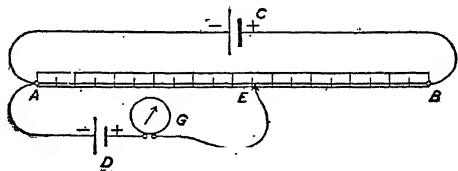
FIG. 358. Potentiometer

be less than that of C, is connected to A and its positive pole is connected through a galvanometer G to the sliding contact E, it will be found on sliding the contact back and forth along the resistance wire that there is a particular position of the contact for which the current in the galvanometer is zero. This can only be true when *the potential of the cell D is exactly balanced by the potential drop along that part of the resistance wire included between its terminals,* marked AE in the figure. Furthermore, if the resistance wire is uniform it is clear that the potential of a cell balanced in this way is proportional to the *length* of the slide wire included between the terminals of the cell, and, therefore, the *ratio of any two voltages balanced in this way is equal to the ratio of the two corresponding lengths intercepted on the slide wire.* For instance, if one cell has its voltage balanced in this way and then this is removed and a second one has its voltage so balanced, the ratio of these two voltages is equal to the ratio of the slide wire readings. By this method the potential of a cell may be accurately determined by comparing it with that of a standard cell whose voltage is exactly known.

652. Ohm's Law Applied to Part of a Circuit. It has been seen (§ 618) that in any whole circuit

$$I = \frac{E}{R} \tag{1}$$

where E is the electromotive force in the circuit and R is a constant known as the resistance of the circuit. Similarly in any *part* of a circuit such as that between A and B (Fig. 359) the current I is proportional to the difference of potential between those points and may be written
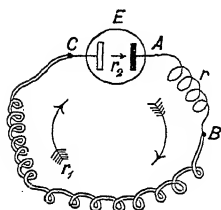
Fig. 359

$$I = \frac{P}{r} \quad \text{or} \quad Ir = P \tag{2}$$

where P is the drop in potential between A and B and r is the resistance of that part of the circuit.

When there is no source of electromotive force, such as a battery cell, in a given portion of a circuit, the difference of potential between its ends in volts is equal to the product of the current in ampères by the resistance of that portion in ohms.

When an electromotive force E is included in any part of the circuit considered, the difference of potential between the ends of that part may be written

$P = Ir - E$ (*E.M.F. in same direction as current*)   (3)

or

$P = Ir + E$ (*E.M.F. acting against the current*).   (4)

For Ir measures the *drop* in potential in the direction of the current, but when E is *with* the current it lifts the potential, as seen in the preceding paragraph. The sign of E is therefore opposite to that of Ir in that case.

653. Resistances in Series. In a complete circuit made up of several conductors, having resistances r, r_1, r_2, and including an electromotive force E, as in figure 359, the successive steps in potential may be written thus

from A to B	resistance $= r$	$P = Ir$	(5)
from B to C	resistance $= r_1$	$P_1 = Ir_1$	(6)
from C to A	resistance $= r_2$	$P_2 = Ir_2 - E$	(7)

The sum $P + P_1 + P_2$ is evidently the total change of potential around the circuit from A around to A again, but this must be zero for it ends at the same potential as it began. Therefore, adding 5, 6, and 7 we have

$$Ir + Ir_1 + Ir_2 - E = 0,$$

but this may be written

$$I = \frac{E}{r + r_1 + r_2},$$

and by Ohm's law $\qquad I = \frac{E}{R},$

hence $R = r + r_1 + r_2$; *i.e., the resistance of several conductors connected in series is the sum of their separate resistances.*

654. Combining Resistances in Parallel. Let three conductors having resistances r_1, r_2, r_3 be joined in parallel in a battery circuit as shown in figure 360. It was shown by Faraday that the sum of the currents in the branches is equal to the total current I before it divides,
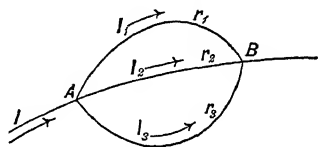
FIG. 360

$$I = I_1 + I_2 + I_3. \tag{1}$$

But the drop in potential from A to B must be the same along either branch. Letting P represent this drop we have

$$I_1 = \frac{P}{r_1} \quad I_2 = \frac{P}{r_2} \quad I_3 = \frac{P}{r_3}.$$

Therefore by (1)

$$I = \frac{P}{r_1} + \frac{P}{r_2} + \frac{P}{r_3} = P\left(\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3}\right).$$

But if R is the effective resistance of the three branches combined

$$I = \frac{P}{R}.$$

Therefore $\qquad \dfrac{1}{R} = \dfrac{1}{r_1} + \dfrac{1}{r_2} + \dfrac{1}{r_3}.$

The reciprocal of the resistance of a conductor is called its *conductance,* hence *the sum of the conductances of several conductors joined in parallel is the conductance of the combination.*

If the three resistances above considered are equal the combination will have one-third the resistance of one alone.

655. Galvanometer Shunt. It frequently happens that a current is to be measured by a galvanometer adapted to smaller currents. In such a case a wire S of suitable resistance, called
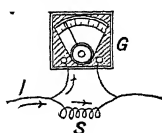
Fig. 361

a *shunt* (Fig. 361), may be connected across from one galvanometer terminal to the other. The current then divides between the shunt and the galvanometer. If the resistance of the galvanometer is just 9 times that of the shunt used, the current will divide in the ratio 1 : 9, so that one-tenth of it flows through the galvanometer and nine-tenths through the shunt.

656. Resistance of Wires and Specific Resistance. The resistance of a wire or of any conductor of uniform cross section increases with its length and is inversely proportional to its cross section. The results of the last two articles show that this is so. For if the cross section of a conductor is doubled it is equivalent to two of the original conductors side by side in parallel, and hence by § 654 the resistance is one-half as much as before.

The resistance of a cylindrical conductor of a given substance one centimeter long and one square centimeter in cross section is called the specific resistance of that substance, or its resistivity.

When a wire of length l and cross section s is made of a substance having *resistivity* ρ, its resistance R is given by the formula

$$R = \frac{\rho l}{s}.$$

657. Resistivities. The curves given in figure 362 show the specific resistances of certain pure metals and alloys and also the variation of the resistances with temperature.

If the curves for the pure metals are produced it will be found that they intersect the base line in the region of the absolute zero $(-273°)$. The experiments of Onnes on the resistance of gold, silver, mercury, lead and tin at very low temperatures show that as the temperature is lowered they approach zero resistance at

a point *a few degrees above the absolute zero*, the change being sudden at the last. The resistance of mercury, for example, decreases slowly from 4.41° to 4.21° above the absolute zero, it then rapidly diminishes and practically disappears at 4.19°. The
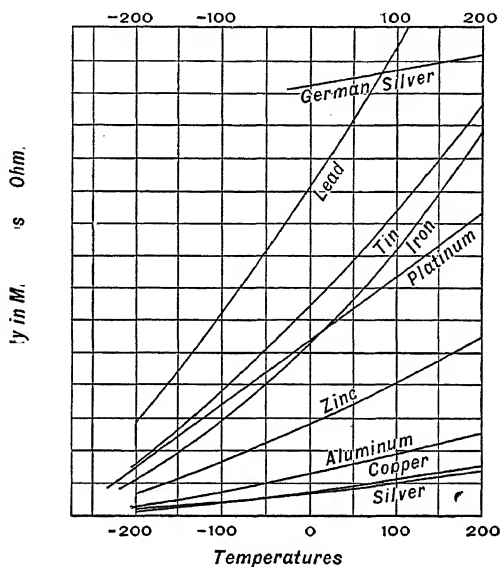
FIG. 362

following table shows some specific resistances at 0° C. in millionths of an ohm with the corresponding increase in resistance per ohm when the temperature is raised from 0° to 100° C. (From Dewar and Fleming.)

SPECIFIC RESISTANCES AT 0° C.

SUBSTANCE	MILLIONTHS OF AN OHM	INCREASE PER OHM FROM 0° TO 100° C.
Platinum...................	10.917	0.367 ohm
Silver.....................	1.468	0.400
Copper....................	1.561	0.428
Iron......................	9.065	0.625
Nickel....................	12.323	0.622
Lead......................	20.380	0.411

The resistance of carbon decreases with rise of temperature instead of increasing, so that the filament of an incandescent lamp may have only one-third as much resistance when hot as when cold.

The resistivities of alloys cannot in general be calculated from those of their constituents, but are often much greater than would be expected. The temperature coefficients of German silver, platinoid, and manganin are much less than those of pure metals; for this reason as well as for their large specific resistances these substances have been used extensively in making resistance coils.

For small ranges of temperatures the resistance, R_t, of a conductor at a temperature $t°$ C. may be expressed by the approximate formula, $R_t = R_0 (1 + \alpha t)$ where R_0 is the resistance at $0°$ C. and α is a constant known as the *temperature coefficient of resistance*. It is the fractional change of resistance for a temperature change of $1°$ C.

Alloys	*Temperature Coefficients*
German-silver (Cu 50, Ni 26, Zn 24). .	0.00040
Platinoid (Cu 60, Ni 14, Zn 24, Tg 2).	0.00022
Manganin (Cu 84, Ni 12, Mn 4)	0.000001

658. Standard Resistance. Standard resistances are made of wire having a small temperature coefficient and not otherwise subject to change. The best coils are made of *manganin*. The coil is provided with heavy copper terminals of almost negligible resistance, and is so mounted that it will quickly take the temperature of the oil bath in which it is immersed, and by which its temperature is maintained constant.
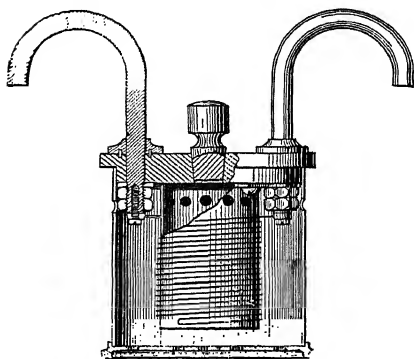
FIG. 363.   Standard of resistance

659. Resistance Boxes. Boxes of coils having different resistances are made so as to be conveniently used in measure-

ments, as shown in figure 364. On the hard-rubber top of the box are mounted a number of blocks of brass which can be connected by brass plugs fitting between them. Within the box are the resistance coils wound on spools, one end of a coil being soldered to one block and the other end to the next one so that one coil bridges each gap. The external circuit is connected at the terminal binding screws, and when all the plugs are *in*, the only resistance is that of the brass blocks and plugs themselves.

FIG. 364. Resistance box

But if a plug is pulled out the current must then flow through the coil joining the blocks, and accordingly that resistance is introduced.
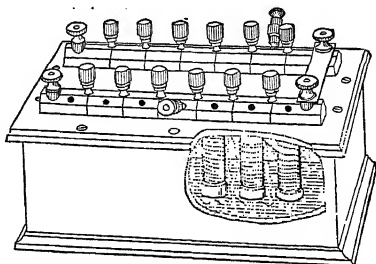
660. Rheostats. Coils of wire so mounted that they can easily be thrown into or out of a circuit to regulate the strength of current without particular reference to measurement are known as *rheostats*. A convenient form is shown in figure 365, where, as the radial arm is moved around the dial from block to block, one coil after another is added to the circuit until as many as may be desired are thrown in.

661. Wheatstone's Bridge. If a current from a battery E (Fig. 366) divides between the two conductors ACB and ADB, the resistances of these branches may be very different and consequently the current in one may be much larger than in the other, but as they both start at the same point A and end together at B, the *fall in potential* must be the same in each, and corresponding to any point, such as C in the one, there must be a point D in the other where the potential is the same.

If p, q, r, s are the resistances of the four segments, AC, CB, AD, DB then it may be shown that $p : q :: r : s$.

Let I_1 be the current in the upper branch and I_2 that in the lower branch, then I_1p is the drop in potential from A to C and is equal to I_2r the drop from A to D, thus

$$I_1p = I_2r, \tag{1}$$

so also
$$I_1q = I_2s \tag{2}$$

and dividing (1) by (2) we find

$$\frac{p}{q} = \frac{r}{s}.$$
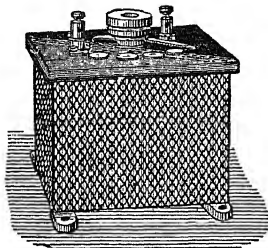
FIG. 365
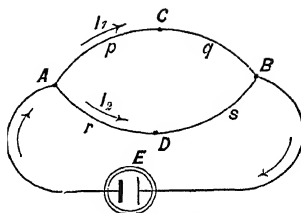
FIG. 366

662. Slide Wire Bridge. The relation just demonstrated is made use of in the comparison of resistances, a convenient device for the purpose being the slide wire bridge shown in figure 367.
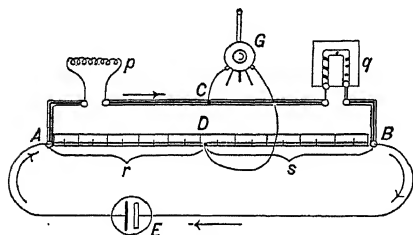
FIG. 367. Slide wire bridge

Suppose p is some coil of wire whose resistance is to be measured by comparison with a standard resistance box q. The current from the battery E divides at A, part flowing through the branch ACB which consists of the two resistances to be compared, p and q, connected by thick copper strips of extremely small resistance, while part flows through the branch ADB which is a uniform wire stretched along a graduated scale.

A sensitive galvanometer has one terminal connected at C midway between p and q while the other is attached to a slider which is moved along the stretched wire until the point D is found where there is *no deflection of the galvanometer*, showing that C and D are at the same potential. Then by the previous paragraph the resistance of p must be to that of q as the resistance of r is to that of s, where r and s are the segments of the bridge wire on each side of D. But since the wire is uniform the *resistances* of r and s are in the same ratio as their *lengths*, so that

the resistance of p is to that of q as the length of r is to the length of s, and therefore p may be calculated by proportion when q is known.

The resistances of the heavy copper connecting strips are so small compared with the resistances of p and q that they may ordinarily be neglected.

663. Resistance Thermometers.   The change in resistance of a coil of wire when its temperature is altered is used for the measurement of temperatures ranging from very low temperatures up to 1500° C.   For small ranges of temperatures the approximate formula $R_t = R_0(1 + \alpha t)$ (see § 657) may be used, and t can be found when R_t, R_0 and α are known.

Callendar was the first to use a coil of fine platinum wire for temperature measurements of extreme precision, obtaining reliable results for changes of temperature as small as one-ten millionth of 1° C.

PROBLEMS

1. An electric car line has a resistance of 0.4 ohm per mile. What is the drop in potential in the line if a car 3 miles from the station is using 50 ampères?

2. If one car 1 mile from the station and another 2 miles from the station are each using 50 ampères, what is the drop in potential to the more distant car, resistance being as in preceding problem?

3. What external resistance when joined to a gravity cell having a resistance of 2 ohms will make the potential difference between the terminals of the cell 0.7 of its electromotive force?

4. A gravity cell of E.M.F. 1 volt and resistance 2 ohms, is connected with another battery by which a current of 1 ampère is made to flow through the cell from the copper to the zinc pole inside the cell. Find the drop in potential in the cell due to resistance, also the difference in potential between the two poles and which is at the higher potential.

5. When the conditions are as in problem 4 except that the current flows through the cell in the opposite direction, find the potential difference between the poles and which is at the higher potential.

6. When the current through the cell of problem 4 is $\frac{1}{2}$ ampère, and flows through the cell from the zinc toward the copper pole, what is the potential difference between the two poles?

7. A gravity cell of resistance 2 ohms and E.M.F. 1 volt, a dry cell having a resistance of 0.5 ohm and E.M.F. 1.4 volts and a wire of resistance 2.3 ohms are joined in series. Find the drop in potential due to resistance in each part of the circuit, also the potential difference between the terminals of each cell.

8. If the cells in the preceding problem are reversed so that one acts against the other, find the drop in potential in each cell and in the external resistance and the potential differences as before.

9. What is the resistance of two conductors connected in parallel, one of 3 and the other of 10 ohms resistance?

10. When a current of 31 ampères divides between three parallel conductors whose resistances are 2, 3, and 5 ohms, respectively, find the current in each branch, also the drop in potential in the parallel combination.

11. What part of the whole current will flow through a galvanometer having a resistance of 5 ohms if shunted by a wire of 0.1 ohm resistance?

Energy and Heating Effect of Current

664. Energy of a Current. When a current flows from a point where the potential is V_1 to another point where it is V_2, each unit charge that passes has less energy at the lower potential than at the higher, and the *difference between the two must be the energy which in some form or other is spent between the two points; it may be in heat in the conductor, or in chemical action, or in doing mechanical work.* When unit charge passes from V_1 to V_2 (Fig. 368) the work expended is $V_1 - V_2$ (§ 571). If Q units pass in t seconds the work is

$$w = Q(V_1 - V_2)$$

and the energy spent per second is

$$\frac{w}{t} = \frac{Q}{t}(V_1 - V_2).$$

But $\frac{Q}{t}$ equals I, the *current*, and the potential difference $V_1 - V_2$ is represented by P.

Therefore
$$\frac{w}{t} = I(V_1 - V_2) = IP.$$

Or, *the energy spent per second in any part of a circuit is the product of the current strength by the fall in potential in that part.*

If the current and potentials are measured in C. G. S. electromagnetic units then the product will give the energy spent in ergs per second. When the current is in ampères and the potentials are in volts the energy per second is given in units called *watts.**

---

* The watt is named in recognition of the researches of James Watt on the power of engines.

Since the volt is 10^8 times the electromagnetic unit of potential, while the ampère is one-tenth the electromagnetic unit of current, it follows that *one watt* = 10^7 *ergs per second*.

Thus a watt represents a certain rate of spending energy per second, it is therefore a unit of power, and bears a definite ratio to other units of power.

1 *Horse-power* = 746 *watts* = 550 *foot-lbs. per second.*

665. Where Energy is Absorbed and Where Given Out. From the diagram (Fig. 368) it is seen that everywhere in the external circuit from C to Z the current flows from points of
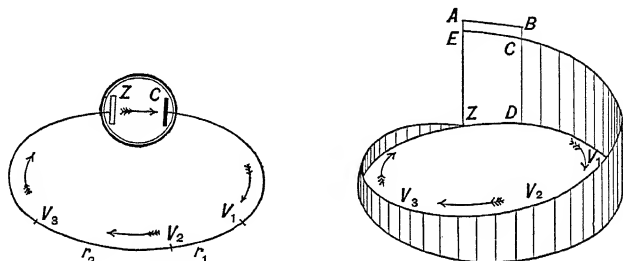
Fig. 368

higher to lower potential, and is therefore spending energy. But there is a great *rise* in potential from the zinc to the acid at A; at this point the current must therefore *receive* energy from the chemical action which effects the transfer in the face of the opposing difference of potential. From A to B there is again a fall in potential and spending of energy within the battery cell, then at B there is a sudden drop of potential which shows that at that point also energy must be spent, but in this case against the chemical forces which at this point exert an electromotive force *against* the current.

The only place where energy is absorbed by the current is at the surface of the zinc plate, and therefore the energy of the chemical action at the zinc plate supplies that which is spent in all other parts of the circuit. This conclusion is based on the law of the conservation of energy.

At any point in the circuit where there is an electromotive force E, the energy taken in or given out per second, is IE, where I

*is the current strength. The energy is absorbed if the current is
with the electromotive force, and given out if the current is against it.*

666. Heating Effect of Current. Suppose a circuit, such as
shown in figure 369, contains a battery B, a coil of wire of
resistance r, a motor M, and a cell N containing some electro-
lyte. Let P, P_1, and P_2 be the poten-
tial differences between the terminals
of the coil, the motor and the cell,
respectively, when there is a current
of I ampères.

Then $IP, IP_1,$ and IP_2 are the watts
spent in the corresponding parts of
the circuit.

In the coil there is no electromotive
force and therefore
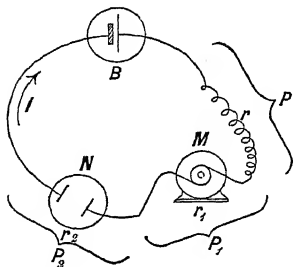
FIG. 369. Composite cir-
cuit including resistance,
motor, and electrolytic cell

$$P = Ir \qquad \text{and} \qquad W = I^2r.$$

*In this case there is no mechanical or chemical work done and
all the energy is spent in heat.*

In the motor there is an electromotive force against the
current as will be shown later (§ 752) and therefore

$$P_1 = Ir_1 + E_1 \qquad \text{and} \qquad W_1 = I^2r_1 + IE_1.$$

The term I^2r_1 represents the power spent in heat while IE_1
is the power spent in driving the motor.

In the third case, where there are chemical changes, there is
usually also an electromotive force which may be either with the
current, as in case of a battery cell, or against it, as in charging a
storage cell; therefore, $P_2 = Ir_2 \pm E_2$

and $$W_2 = I^2r_2 \pm IE_2.$$

Here, again, I^2r_2 is the watts spent in heat, while IE_2 is the
watts spent in chemical work or received from chemical work
as the case may be. (In which case is the sign to be taken
plus?)

The *heating effect* of a current of I ampères in a resistance of
r ohms may always be expressed in watts by the formula

$$W = I^2r.$$

But 1 watt $= 10^7$ ergs per second, and 1 gram-calorie of heat is equivalent to 4.187×10^7 ergs, therefore, 4.187 watts $= 1$ gram-calorie of heat per second, and we have,

$$heat\ in\ gram\text{-}calories\ per\ sec. = \frac{watts}{4.187}.$$

Summary

The *total* watts spent in any portion of a circuit, in which P is the fall in potential in volts and I is the current in ampères, is given by the formula

$$W = IP,$$

the watts spent in *heat* by

$$W = I^2r,$$

while the heat in gram-calories per second is expressed by

$$H = \frac{I^2r}{4.187}.$$

667. Electrical Calorimeter. The heat developed in a conductor may be readily measured by a calorimeter such as shown in the figure. A coil of wire is immersed in a non-conducting liquid (distilled water may be used) contained in a calorimeter which is screened from outside radiation. The current passing through the coil is measured, and also the difference in potential of the two terminals, and so the electrical work can be calculated and compared with the heat developed, which is determined by the rise in temperature of the liquid in the calorimeter when its mass and specific heat are known.

668. Incandescent Electric Lamps. In the ordinary incandescent lamp used in electric lighting the current passes through a fine filament of carbon or tungsten enclosed in a glass bulb from which the air is thoroughly exhausted. The ends of the filament are joined to wires which must be of a special material where they are sealed into the glass. This material must have the same
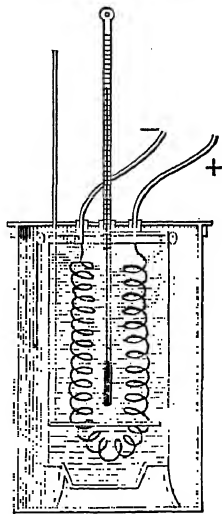
Fig. 370. Electrical calorimeter

coefficient of expansion as glass so the joint will not crack with changes of temperature. For many years platinum was used for this purpose but its use has been replaced by *platinum substitute* which is an alloy of iron and nickel. The essential requirements for the filament are strength and ability to withstand high temperature for a reasonable length of time. In the original incandescent lamp invented by Edison a carbon filament was used. Ordinary carbon lamps require about 3 watts per candle-power and operate at a filament temperature of about 1900° C., and though by raising the temperature of the filament the candle-power per watt is greatly increased, the filament rapidly volatilizes, blackening the bulb, and finally breaks.

The carbon filament lamp has been almost entirely replaced by one with a tungsten filament because of its higher efficiency. This metal which has the very high melting point of 3400° C. was first produced in ductile form suitable for drawing into filaments about the year 1910. The tungsten filament in vacuum is ordinarily operated at a temperature of about 2200° C. and requires about 1.35 watts per candle-power, but if its temperature is raised to 2800° C. it requires less than 0.5 watt per candle-power in vacuum, though it volatilizes so rapidly that the life of the lamp is short. When the bulb is filled with nitrogen or argon, however, the evaporation of the filament is only about 0.01 of what it is in vacuum, and so the lamp may be run profitably at the higher temperature. Such a nitrogen filled lamp may operate at a temperature of 2800° C. requiring something over 0.5 watt per candle-power. Filling the bulb with gas causes some loss of luminous efficiency due to conduction of heat away from the filament by the gas. The filament is, therefore, made of a rather closely wound helix, for the cooling effect of the gas is then much less than when the turns of the filament are widely separated. The most efficient incandescent lamp used is the moving picture gas-filled lamp which operates at a temperature of about 3000° C. with a power consumption of only about 0.45 watt per candle-power.

Incandescent lamps are usually connected in parallel between two conductors as A and B (Fig. 371) which are maintained at a constant difference of potential by a dynamo or battery of *low resistance*. They are called 50-volt lamps when they are brought to their proper luminosity by a difference of potential of 50 volts between their terminals.

The candle-power depends on the power supplied to the lamp and therefore both on the current and voltage. Good tungsten lamps as ordinarily used have an efficiency of about 1.35 watts per candle-power, so that a 50-watt lamp gives about 36 candle-power.

Comparison of a 50-volt and 100-volt 50-watt lamp, is shown in the table:

Voltage	Current	Resistance	Power
50 volts	1 ampère	50 ohms	50 watts
100 volts	$\frac{1}{2}$ ampère	200 ohms	50 watts

As the current in the 100-volt lamp is only one-half as great as in the 50-volt lamp, there is only one-fourth as much loss in heat in wires of a given resistance leading to the lamps.

669. Incandescent Lamps in Series. The mode of connecting incandescent lamps in parallel shown in figure 371 has the advantage that any one lamp can be turned on or off without particularly disturbing the others. But the current in the main wires is the sum of the currents in the lamps and, as the energy spent in the circuit is proportional to the *square* of the current, the loss of energy in the main wires will be serious unless they are large and of low resistance. For street lighting, lamps are commonly connected in series so that, however many lights there may be, the current in the conducting wires is no greater than for one lamp. There are two considerations which prevent this system from being used in house lighting. First, the potentials required are dangerous. If 30 lamps are connected in series and each requires 20 volts, the dynamo must have an electromotive

FIG. 371.   Lamps in parallel   FIG. 372.   Lamps in series

force of 600 volts, and if a break or interruption of the circuit occurs at any point, the full difference of 600 volts will be experienced there.   Second, if the filament of one lamp breaks it stops the current and all the lamps go out. This latter difficulty is overcome most ingeniously in street lighting by arranging a little side circuit or *by pass* in each lamp which is complete except at one point where a slip of paper is interposed.   When the lamp is acting the current passes wholly through the filament; but if this breaks, the current is interrupted and immediately the whole electromotive force of the dynamo is brought to bear on the paper which is thereby punctured, permitting the current to pass through the side circuit.

In street lighting large gas-filled tungsten filament lamps are used more than any other form of lamp because of their high efficiency and reliability.

670. Electric Arc. In the year 1801, Sir Humphrey Davy, who had constructed an immense battery of 1000 cells, observed that when the terminal wires were touched together for an instant and then drawn apart the discharge took place through the air like a stream of fire from one pole to the other, and at the same time the tips of the wires were intensely heated. The effect was most marked when carbon rods were used for the terminals. When the discharge took place horizontally it was bent upward like a bow (on account of the heated air rising) and so Davy called it the *arc* discharge. This tendency to curve out to one side is noticed in every long arc whatever its position. In arc lamps the carbons are only slightly separated, both tips are intensely heated, particles of carbon are carried across from

the positive to the negative carbon causing a crater-like cup on the end of the positive carbon while the negative carbon is pointed; the positive carbon also is used up about twice as fast as the negative. The point of most intense luminosity is in the crater of the positive carbon where the temperature is found to be about 3500° C. The difference in luminous power of large and small carbon arcs seems to be due to the greater extent of luminous surface in one than in the other, the actual brightness of the glowing surface being the same in all.

An electromotive force of 40 volts is required to maintain an arc between carbons. The temperature of the arc is the highest that has been produced by artificial means. Copper, iron, gold, silver, and platinum, if placed in it, are melted and volatilized.

671. Street Arc Lamps. Carbon arc lamps were widely used at one time in street lighting because of their comparatively high efficiency. The tungsten filament gas-filled lamp has a still higher efficiency, and its only competitor in street lighting is the *magnetite* arc lamp, whose efficiency is about the same, that is, it consumes about 0.5 watt per candle-power. The so-called *flaming* arc has almost as good an efficiency and is sometimes used because of its brilliant orange light. The magnetite arc has the positive electrode made of copper and the negative made of magnetite or black oxide of iron mixed with oxides of titanium and chromium. The positive copper electrode is the upper one with the negative magnetite electrode below. The magnetite electrode alone is consumed, during which process it is further oxidized to the higher red oxide of iron. The flaming arc uses carbon electrodes with a core chiefly of calcium fluoride. Both of these types of arc differ from the ordinary carbon arc in that the light emitted comes largely from the luminous vapor between the electrodes, rather than from the electrodes themselves, which are much the brightest part of the carbon arc. Street arc lamps are provided with a special magnetic regulating device which keeps the electrodes just the right distance apart during operation.

Arc lamps are connected in series for street lighting because a current of only 4 or 5 ampères is then needed; but as 75 volts must be allowed for each lamp, to operate 40 lamps on the one circuit an electromotive force of 3000 volts is required; hence arc light circuits are dangerous. Lamps connected in this way have a device by which if the electrodes become caught and do not make the arc at all the current can still flow through the lamp.

672. Search Lights. For large search lights the carbon arc is used and is so mounted that almost a point source of light is obtained from the intensely hot spot in the crater of the positive carbon. This makes it possible to produce a very concentrated beam when a reflector is used to direct the

light. Although the magnetite arc is more efficient, its light comes from a considerable area of luminous vapor distributed between the electrodes and the light therefore cannot be directed by a reflector without a scattering and a consequent loss of concentration.

673. Electric Furnace. In the production of aluminum, carborundum, calcium carbide, high grade steel and many other materials, electric furnaces are employed. In these furnaces great carbons 2 feet in diameter are mounted and imbedded in the materials that are to be heated, the whole is surrounded by walls of brick or fire clay, and the electric arc is established between the carbons. Under the combined influence of the enormous heat and the electrolytic action of the current the desired transformations are wrought.

PROBLEMS

1. A current of 14 ampères divides between two branches, one of 2 ohms and one of 5 ohms resistance. Find the current in each branch, and the watts spent in each. In which resistance is the greater amount of heat developed per second?

2. The terminals of a gravity cell of 2 ohms resistance and 1 volt E.M.F. are connected with a coil of resistance 3 ohms. Find the watts spent in heat in the coil and also in the cell, also the total watts supplied by the cell.

3. What must be the resistance of a coil of wire in order that a current of 2 ampères flowing through the coil may give out 1200 gram-calories of heat per minute?

4. If the difference in potential of the ends of a coil is 50 volts, what must be its resistance that 500 gram-calories of heat may be developed in it per second?

5. Find the gram-calories per second developed in each of two coils; one having resistance 3 ohms and current 6 ampères, the other a resistance of 4 ohms and a difference of potential of 20 volts between its ends.

6. How many horse-power must be expended to maintain 200 100-volt lamps in operation, each lamp taking $\frac{1}{2}$ ampère of current and having a potential difference of 100 volts between its terminals?

7. How many horse-power are required to operate a series of 60 incandescent street lamps in series, the current in each lamp being 3 ampères and the resistance per lamp being 7 ohms?

8. In an electric railway having a total line resistance of 0.4 ohm per mile, what is the loss in horse-power in two miles of line when a current of 50 ampères is being supplied to a distant car?

9. At 10 cents a kilowatt-hour what is the cost of heating 1000 liters or a cubic meter of water from 20° C. up to 90° C. by electricity?

674. Seebeck's Discovery. In 1821 Seebeck, of Berlin, discovered that *in a circuit made of two different metals if one junction is hotter than the other there is an electromotive force which causes an electric current.* This electromotive force is generally very small compared with ordinary battery cells, and consequently to obtain much current the circuit must have very low
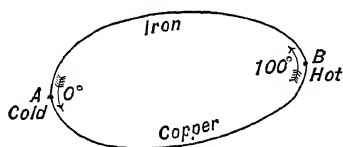
FIG. 373
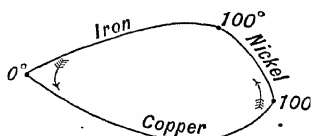
FIG. 374

resistance. For example, in the copper-iron circuit shown in figure 373 when one junction is at 100° and the other at 0°, the electromotive force is about 0.001 of a volt and causes an electric current from copper to iron at the hot junction and from iron to copper at the cold one. *The introduction of another metal does not make any difference provided the two junctions of the new metal are at the same temperature.*

For example, the electromotive force is the same in the three circuits shown in figures 373 and 374.

675. Thermopile. In order to obtain larger electromotive forces pairs of metals are combined in series to form *thermopiles.* The form devised by Nobili and used by Melloni in his researches on heat radiation consists of alternate strips of antimony and bismuth connected as shown in the figure, and carefully insulated from each other except at the junctions, where they are soldered together. These metals were chosen because they give a large

FIG. 375. Thermopile diagram

electromotive force which acts from bismuth to antimony at the hot junctions and from antimony to bismuth at the cold.

Rubens has improved the thermopile by using fine wires of iron and constantan (a nickel alloy) in place of antimony and bismuth. The mass to be heated in this case is very small so that it warms quickly when exposed to radiation.

The thermopile is usually mounted in a metal case so that only one set of ends is exposed to the source of heat to be investigated. If its terminals are connected to a sensitive galvanometer of low resistance, it becomes an exceedingly delicate means of measuring heat radiation.

676. Change of Thermoelectric Force with Temperature. If one junction of a copper-iron circuit is kept at 0° C. while the other is steadily raised in temperature, the electromotive force is found to increase rapidly at first, then more gradually, reaching a maximum when the hot junction is at 260° C., after which the electromotive force falls off, becoming zero at 520° C. If the junction is heated still hotter the electromotive force reverses and the current flows from iron to copper at the hot junction. If the observations are plotted with the temperatures of the hot junction as abscissas and electromotive forces as ordinates a curve such as shown in figure 376 is obtained. It is a parabola and is perfectly symmetrical about the vertical line through its vertex, which corresponds to the temperature of maximum electromotive force.
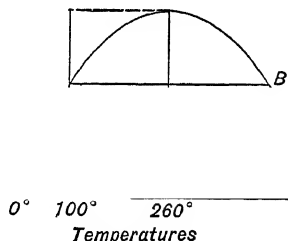
Fig. 376. Thermoelectric curve of e. m. f. of copper and iron

This reversal of the thermoelectric current was discovered by Cumming in 1823.

If the cold junction is kept at 100° instead of zero the curve will be exactly the same except that the origin of coördinates will be moved from O to A, and electromotive forces will now be measured from the base line AB.

677. Thermoelectric Powers. It is clear from the foregoing that the inclination of the curve at any point, or *the rate of*

change of electromotive force per degree change in temperature, depends only on the temperature of the junction which is being warmed or cooled and not at all on the temperature of the other junction, provided it is constant.

This change of electromotive force per degree change of temperature of a junction is known as the relative thermoelectric power of the substances involved.

If the thermoelectric powers of iron and copper are plotted as ordinates along a scale of temperatures we shall obtain the diagram shown in figure 377. The curve is a straight line intersecting the axes at 260° C.; for at that temperature the relative thermoelectric power is zero, as is seen also from the curve in figure 376, where the maximum point is at 260°, showing that a
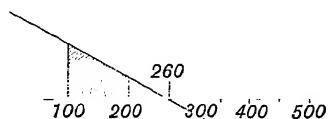
FIG. 377. Thermoelectric powers of iron and copper

small change in temperature produces *no* change in electromotive force. This is called the *neutral temperature* for these two metals.

678. Thermoelectric Diagram. In the thermoelectric diagram devised by Tait, the thermoelectric powers of the metals referred to lead are plotted as ordinates along a scale of temperatures; lead being taken as standard because in it the Thomson effect (§ 681) is zero. Such a diagram is shown in figure 378. It will be observed that within the limits of the diagram the variations with temperature of the thermoelectric powers of the metals are represented by straight lines.

The electromotive force of a couple made of any two metals is expressed by the area included between the lines of the two metals and the ordinates of the temperatures of the junctions.

The diagram is so constructed that the *direction* of the resultant electromotive force is *clockwise;* that is, in case of iron and copper between 0° and 100° the current will be from copper to iron at the hot junction.

679. Peltier Effect. It was discovered by Peltier in 1834 that if a current of electricity flows around a circuit made up of two metals heat will be given out at one junction and absorbed at the other.

A beautiful demonstration of the Peltier effect was given by Tyndall by means of an ordinary thermopile. A thermopile is taken in which all parts are at the same temperature, so that it gives no current. On connecting it for a few seconds to a battery and then disconnecting it and joining it to a galvanometer a decided current is observed, showing that one set of junctions must have been more heated by the current than the other set. The current obtained is opposite to the first and tends to restore the equality of temperature disturbed by the first, for one stored up heat energy in the thermopile and the other transforms that energy back again into energy of current.
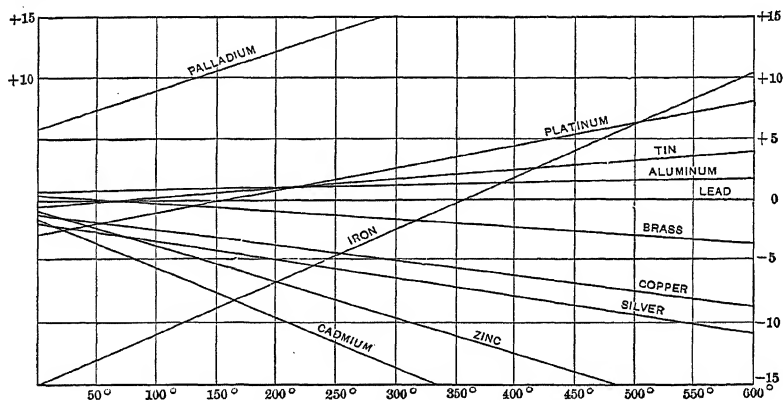
FIG. 378. Thermoelectric diagram

Thermoelectric powers are given in micro-volts per degree

By a thermopile there is a direct transformation of heat energy into electrical energy, but it is not efficient because there is a serious loss of heat by conduction from the hot junctions to the cold.

680. The Conservation of Energy in Thermoelectricity. The Peltier effect affords a beautiful illustration of the principle that energy is absorbed at those points in a circuit where there is an electromotive force acting *with* the current and is given out at those points where there is an electromotive force acting *against* the current (§ 665).

In a circuit of two metals all at one temperature there may be electromotive forces at the two junctions, but since the

temperature is the same at both, these electromotive forces are equal and opposite and consequently there is no current. If a current is now caused to flow by means of a battery, energy is given out at the junction where the electromotive force is against the current and that junction is heated, while the other is cooled. The two junctions no longer balance each other, and it is clear that *the resultant electromotive force* which arises from the change in temperature *must be against the current which brought it about.* Otherwise in a simple closed circuit of two metals if one junction were heated a little to begin with, a current would be set up which would still further increase the difference in temperature of the junctions and would so become continually stronger and might be used to run a motor and do mechanical work until all the heat energy in the thermopile was used up and it was reduced to the absolute zero of temperature.

681. Thomson Effect. In 1854, Lord Kelvin (Sir William Thomson) showed that in a thermoelectric circuit there must in general be electromotive forces not only at the junctions, but also in the homogeneous conductors between the junctions, as they are not at the same temperature throughout.

This effect was predicted by Lord Kelvin as a consequence of the law of energy and was then verified by the following experiment.
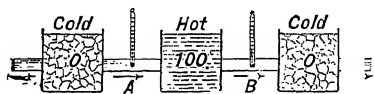
FIG. 379

A bar of iron was set up as shown in figure 379 so that the center was heated by boiling water while the ends were cooled with ice. When a current was established all parts of the bar were warmed, but a thermometer at *A* was observed to stand higher when the current was from left to right than when it was reversed, while the opposite was true at *B*.

682. Applications. The *thermopile* as a delicate means of observing the intensity of heat radiation has already been described (§ 675).

A particularly sensitive instrument for the same purpose was devised by Boys and is known as the *radio-micrometer*. In this instrument a simple circuit of bismuth and antimony is suspended between the poles of a powerful magnet by a fine quartz fiber. One of the two junctions is protected from outside radiation by the surrounding instrument, while the other hangs in an opening so that radiation may be directed upon it. The slightest difference of temperature causes an electromotive force and since the resistance

of so short a circuit is very small a comparatively large current is produced, which, reacting on the magnetic field, causes the suspended circuit to turn. A light mirror mounted on the suspended system turns with it so that the angular deflection may be read by a telescope and scale.

For the measurement of *high temperatures* a thermocouple consisting of a wire of pure platinum joined to another of an alloy of platinum and rhodium may be used. In the Le Chatelier pyrometer such a couple, mounted in a protecting sheath of porcelain, is thrust into the furnace or oven of which the temperature is to be determined; wires from the couple lead to a suitable galvanometer graduated to read temperatures directly up to 1500° C.

For the measurement of ordinary temperatures a thermocouple of iron and German silver is often convenient.

PROBLEMS

1. Find the thermal electromotive force of an iron-copper circuit in which one junction is at 0° and the other at 200° C.

2. Find the increase in electromotive force in a lead-iron circuit when the temperature of the hot junction is changed from 150° to 151°.

3. What relation must the lines of two metals on the thermoelectric diagram bear to each other in order that the increase in electromotive force per degree rise in temperature of the hot junction may be a constant?

4. When one junction of zinc-iron circuit is at 50° C., at what different temperature may the other junction be without causing any current in the circuit?

MAGNETIC EFFECTS OF CURRENTS

683. Oersted's Experiment. The first evidence of the magnetic action of an electric current was obtained in 1819 by the Danish physicist, Oersted, who discovered that when a wire carrying a current is held in a north and south direction over or under a balanced magnetic needle the needle is deflected as shown in figure 380; and if the directive force of the earth's magnetism is neutralized by means of a magnet, the needle sets itself at right angles to the current.

Wire above needle   Wire under needle
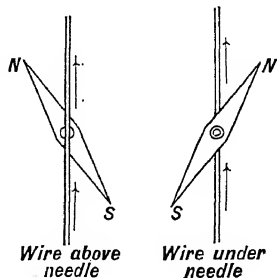
FIG. 380. Oersted's experiment

684. Magnetic Field Around a Straight Conductor. The experiment of Oersted indicates that the magnetic force due to a current is in a plane at right angles to the current. To in-

vestigate its direction more fully cause a strong current to flow in a wire which passes vertically through a card on which some fine iron filings are scattered; on tapping the card the filings arrange themselves in circles about the wire as shown in figure 381. If the current is *down* as shown by the arrows, a small compass needle near the wire, at any point such as P, will point with its north pole in the direction of the arrow at that point, tangent to the circle. If the current is reversed the compass needle will point in the opposite direction.

The lines of magnetic force about a straight conductor carrying a current are circles of which the conductor is the axis.
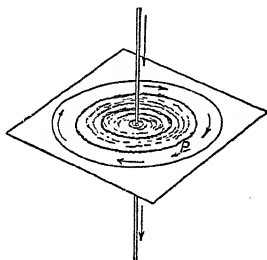
FIG. 381. Field around current   FIG. 382. Right-handed screw

By a comparison of figures 381 and 382 it will be seen that *the positive direction of the lines of force bears the same relation to the direction of the current as the direction of rotation of a right-handed corkscrew bears to the direction in which it advances.*

Another rule that may be given is that *if an observer looks along a conductor in the positive direction of the current, the positive direction of the lines of force as he sees them is clockwise.*

685. Strength of the Field. The strength of the magnetic field near a straight conductor is greatest next to the conductor and diminishes as the distance increases.

The strength of field H, at a distance r from the axis of a long straight wire carrying a current of strength I, is given by the expression

$$H = \frac{2I}{r} \qquad \text{(all quantities in C. G. S. electromagnetic units)}$$

or $\qquad H = \dfrac{2I}{10r} \qquad$ (I in ampères, H in C. G. S. electromagnetic units)

since the ampère is one-tenth the C. G. S. electromagnetic unit of current (§ 616). This formula assumes that the return circuit is so far off that its magnetic effect at the point considered may be neglected.

As the card is tapped on which the iron filings rest, in the experiment described in the last article, the filings work toward the center, the circles gradually getting smaller, for the filings are drawn toward the stronger part of the field.

If a fine copper wire carrying a rather strong current is dipped into some fine iron filings they will cling together in little circular filaments, forming a mass around the wire.

§86. Field of a Circular Current. When a conductor carrying a current is bent into a circle the lines of force are crowded together within the circle and spread out outside. In this case, shown in figure 383, all parts of the circuit conspire to cause magnetic lines of force of which the direction is through the circuit perpendicular to its plane on the inside, and back again on the outside, as shown in the diagram. The lines of force very near the wire are nearly circles about the wire, while at the center they are nearly straight and perpendicular to the plane of the coil.
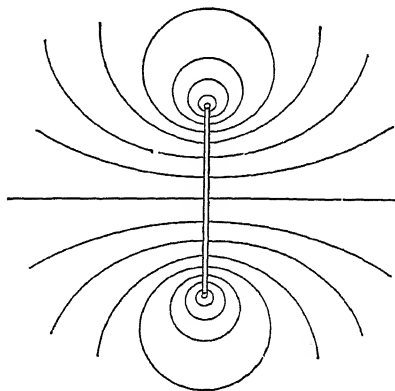
FIG. 383. Field of circular current

If the wire carrying the current makes two turns around the circle instead of one, the magnetic force will everywhere be doubled, and so on for any number of turns.

The strength of the magnetic field at the center of a circular current is proportional to the total length of the conductor wound in the circle and to the strength of the current and inversely proportional to the square of the distance of the conductor from the center.

Thus if r is the radius of the coil and if n is the total number of turns, the length of the wire in the coil is $2\pi r n$, and when the current I is measured in C. G. S. units as defined in § 616, the

strength of the magnetic field H at the center in C. G. S. units is given by the formula:

$$H = \frac{2\pi r n I}{r^2} \quad \frac{2\pi n I}{r} \quad (I \text{ in C. G. S. units})$$

or, since the ampère is one-tenth the C. G. S. unit current,

$$H = \frac{2\pi n I}{10 r} \quad (I \text{ measured in ampères}).$$

The formula assumes that the cross section of the coil is negligibly small compared with r.

By measuring the magnetic force at the center of a coil of known dimensions, the number of ampères of current may be determined.

687. Rowland's Discovery. Rowland discovered that a disc of ebonite, charged with electricity and rotating at high speed, acted upon a magnetic needle placed near it, just as a circular current would. The magnetic effect was found to be proportional to the speed of the disc. This remarkable experiment was carried out by him in 1875.
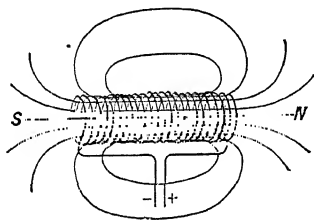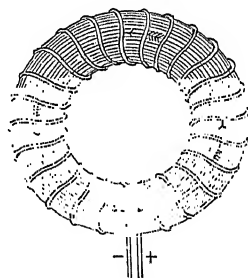
FIG. 384.   Solenoid   FIG. 385.   Ring solenoid

688. Solenoid. A long helix, such as shown in figure 384, is known as a solenoid, and may be wound with one or several layers. When a current passes through such a coil all the turns act together to cause a field of magnetic force in which lines of force pass lengthwise through the interior looping back around the outside. *Looking through the solenoid in the positive direction of the lines of force, the direction of the current is clockwise.* Inside the solenoid a strong magnetic field of great uniformity is produced.

The system of lines of force of a solenoid is thus like that of a bar magnet, the south pole corresponding to that end of the solenoid about which the current flows clockwise, as seen by an observer facing that end.

If such a solenoid is mounted so that it can turn, it behaves like a suspended bar magnet when another solenoid or a bar magnet is presented to it.

689. Iron in a Solenoid. *If a soft-iron core is introduced into a solenoid the number of lines of force is greatly increased,* it may be several hundred times, so that it acts as a strong magnet. A hard-steel core does not have so great an effect in increasing the number of lines of force, though it largely retains its magnetism after the current stops.

The precise effect in such a case depends on the relative proportions of the solenoid and its core. In a short broad solenoid an iron core which fills it will not increase the number of lines of force so much as if the core and solenoid were longer in proportion, for the strong poles exert a magnetic force which in the interior of the solenoid is opposite to that of the coil, as explained in § 513.

690. Ring Solenoid. If a long solenoid is bent into a ring so that its two ends come together as shown in figure 385, a ring solenoid is obtained, and when a current flows in such a solenoid it produces a very nearly uniform field in the interior, though, since the lines of force are not straight but circles, the force must really be slightly stronger toward the inside where the lines of force are smaller circles. *There is no magnetic force outside of such a solenoid.*

If the interior of the solenoid is filled with a ring of iron, all parts of the iron experience the same magnetizing force and there are no poles to complicate matters, so that the *permeability* (§ 516) of the iron can be immediately determined from the increase in the number of lines of force due to its presence. If, for example, the total number of lines of force in the iron is 1000 times what it would have been in the same space if the iron had not been there, the permeability of the iron is said to be 1000.

It is easy to measure by electromagnetic induction (§ 720) the changes that take place in the number of lines of force through the ring and in this way the permeability of iron was studied by Rowland.

691. Magnetic Induction or Flux Density, Permeability and Magnetizing Force. The strength of the field inside a ring solenoid when no iron is present may be called the magnetizing force. Let it be represented by H which will thus express *the number of lines of force per square centimeter of cross section* before the iron is introduced.

The number of lines of force in the iron core per square centimeter of section is called its magnetic induction or flux density and may be represented by B (see § 512).

We then have

$$\mu = \frac{B}{H} \quad \text{or} \quad B = \mu H$$

where μ represents the permeability of the iron.

The relation between the flux density B in iron and the magnetizing force H as the latter is increased, starting at zero, is shown in the curve ab of figure 386, in which abscissas represent values of H and ordinates the corresponding values of B. From this curve it appears that at first B increases slowly, then rapidly, and finally at b as the iron approaches what is called *saturation* a considerable increase in H causes only a small increase in B.

The curve shows that the permeability of iron is not a constant but increases with increase in magnetizing force up to a maximum, after which it rapidly diminishes. Some values are given below.

PERMEABILITY OF A SAMPLE OF SOFT IRON

H	B	μ	H	B	μ
1	1,000	1,000	4	9,700	2,425
2	6,000	3,000	5	11,800	1,966
3	8,200	2,733	17	13,000	765

692. Hysteresis. The changes which take place in iron when its magnetism is *reversed* were first thoroughly studied by Professor Ewing of Cambridge University. The curve of figure 386 shows the changes in induction in soft iron when the magnetizing force is changed from $+13$ to -13 and back again. The rise of induction when the iron is first magnetized is shown by the curve from a to b. The flux density is here 13,000 while H is 13.

On reducing the magnetizing force to zero the induction falls only to 10,800 following the curve bc. By means of a gradually increasing reversed current the magnetizing force is made negative until when $H = -2.0$ the flux density is zero and there are no lines of force in the iron. As the force is made still more negative the iron becomes oppositely magnetized, reaching the value $B = -13,000$ when $H = -13$. Then as the force is again reduced to zero the induction drops to $-10,800$ following the lower curve, and does not become zero till H is $+2.0$. If
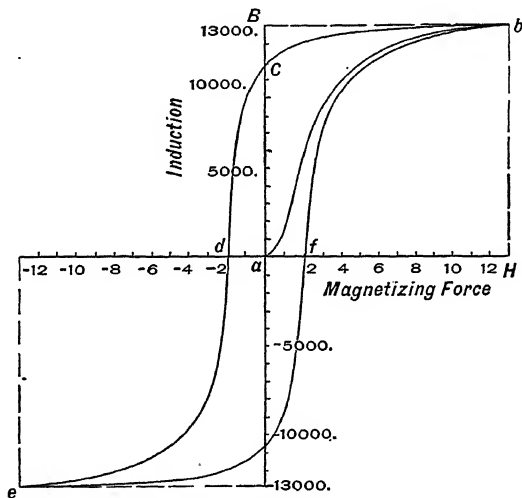
Fig. 386. Hysteresis curve

the magnetizing force is carried up to the former maximum and then again diminished the curve rises to b and then falls back to c exactly as before.

This *lag* of the induction in iron and steel behind the magnetizing force was named by Ewing *hysteresis*. In consequence of it, when a mass of iron is put through such a complete cycle of changes, more energy is spent in magnetizing than is given back when it is demagnetized, the difference being a certain amount lost in heat. The amount lost in this way per cubic centimeter of iron is proportional to the area of the loop of the hysteresis curve, and with a maximum induction of 5000 it may

amount to as much as 2500 ergs per cubic centimeter in each cycle of magnetic change.

Every time the magnetism in a mass of iron is reversed it is put through such changes and since in the iron cores of transformers and dynamo armatures the reversals take place many times in a second it is important in such cases that soft iron should be used in which the hysteresis loss is small. The loss of energy due to this cause may amount to 1 per cent in a transformer made of fairly good iron.

693. Electromagnets. Powerful magnets are made by surrounding soft iron cores with magnetizing coils, as was first shown

FIG. 387. Electromagnet

by the French physicist, Arago, in 1820. A typical form of electromagnet is shown in figure 387.
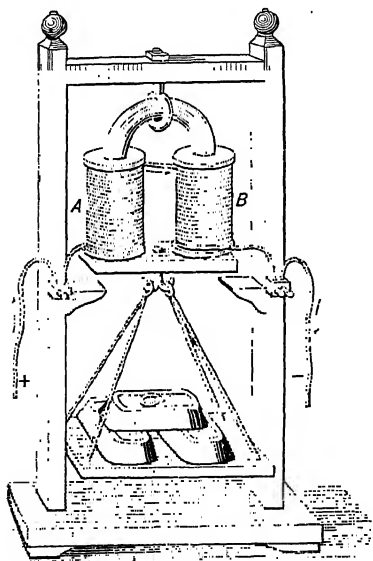
On each of the two arms of a *U*-shaped piece of iron is fitted a cylindrical coil made of wire wound with silk or cotton insulation to separate one turn from another. The coils are so connected that the current flows in opposite directions around the two legs of the magnet, making one end a north pole and the other a south pole. When the soft-iron armature is placed across the two poles a closed circuit of iron is formed so that the magnet with its armature resembles somewhat the ring solenoid with iron core described in § 690. If the armature is sufficiently large most of the magnetic flux will be in the iron, the lines of force being closed curves. The whole number of lines of force established in the core of an electromagnet may be considered as due to the relation of two factors, *the magnetizing power of the current in the magnet coils*, called *the magnetomotive force*, and *the resistance to magnetization offered by the iron core*, called *its reluctance*.

$$Flux = Number\ of\ lines\ of\ force = \frac{Magnetomotive\ force}{Reluctance\ of\ core}.$$

In case of a uniformly wound ring solenoid the magnetomotive force may be shown to be $4\pi nI$ where n is the number of turns of wire around the core and I is the magnetizing current, and we have

$$N = \frac{4\pi nI}{R}$$

where N is the number of lines of force through the core and R is its reluctance, all the quantities being in C. G. S. electromagnetic units.

The above formula applies also approximately to ordinary electromagnets having nearly continuous iron cores.

If the current I' is given in ampères while R and N are in C. G. S. electromagnetic units the formula becomes

$$N = \frac{4\pi nI'}{10R}$$

since the C. G. S. electromagnetic unit of current is equal to 10 ampères.

The product nI' is called the ampère-turns and *the strength of a magnet excited by a small current making many turns is the same as with a large current making few turns, provided the ampère-turns are the same* in both cases.

The shorter the iron circuit and the greater its cross section the less will be the reluctance and the more lines of force will be established by a given number of ampère-turns.

The reluctance of an iron ring may be calculated from the formula

$$R = \frac{l}{\mu A}$$

where l is the mean length of the ring, A is its cross section, and μ is the permeability of the iron. If a circuit is made up of parts that have different permeabilities their reluctances must be calculated separately and added together when the parts are in series.

When the armature is not across the poles the reluctance is greatly increased because of the small permeability of the air

through which the lines of force must pass. Therefore the flux or the number of lines of force established in that case is very much less than with the armature across the poles.

The force with which such a magnet holds its armature is proportional to the area of its poles, and is expressed by the approximate formula

$$\text{Attraction in dynes} = \frac{B^2 A}{8\pi}$$

where B represents the induction or number of lines of force per square centimeter of the surface between pole and armature, and A is the combined areas of the two poles.

694. The Electric and Magnetic Circuits. An interesting parallelism exists between the phenomena of the electric and magnetic circuits which is a useful aid to the memory. Both may be looked at from the point of view of the electric current circuit. The analogy is best brought out by the following table.

Electric Current Circuit	Magnetic Circuit
Applied potential difference $= V$	Applied magnetomotive force $= M$
Resistance $$= R = \frac{\rho l}{s} \quad (\S\ 656)$$ $$= \frac{1}{conductance}$$	Reluctance $$= R = \frac{l}{\mu A} \quad (\S\ 693)$$ $$= \frac{1}{permanence}$$
Current $$I = \frac{potential\ difference}{resistance}$$ $$= \frac{Vcs}{l} \quad (\S\ 618\ \text{and}\ \S\ 652)$$	Magnetic flux $$N = \frac{magnetomotive\ force}{reluctance}$$ $$= \frac{M\mu A}{l} \quad (\S\ 693)$$
Symbols l = length of electric circuit s = area of cross section of path c = conductivity of material of circuit $= \dfrac{1}{\rho}$ (§ 656)	Symbols l = length of magnetic circuit A = area of cross section of path of flux μ = permeability of material of circuit (§ 691)

PROBLEMS

1. What is the strength of the magnetic field 15 cms. from a straight wire carrying a current of 6 ampères?

2. A wire 3 meters long is made into a circular coil with a mean radius of 6 cms. Find the strength of field produced at the center of the coil by a current of 0.1 ampère in the wire.

3. How much current is flowing in one rail of an electric railway which runs in a north and south direction and causes a deflection of 45° in a compass needle held 30 cms. above the center of the rail; taking the strength of the horizontal component of the earth's magnetic field as 0.20?

4. Find the strength of the magnetic field at the center of a circular coil of 7 turns of wire 18 cms. in diameter when carrying a current of 3 ampères.

5. What will be the deflection of a magnetic needle at the center of the coil in the last problem if the coil is placed with its plane vertical and in the magnetic meridian at a point where the earth's horizontal force is 0.16?

6. A ring solenoid has a cross section of 9 sq. cms. There are 8 turns of wire per cm. of length of the solenoid and its whole length is 30 cms. measured along its axis. What magnetomotive force is produced by a current of 1 ampère in the coil? How many lines of force will there be in the solenoid if it does not contain iron? And how many if filled with an iron core of permeability 100?

7. How many lines of force will be set up in a horseshoe magnet with iron armature, the iron circuit having an average cross section of 36 sq. cms., each leg being 15 cms. long and the two legs 12 cms. apart between centers? On each leg is a coil of 400 turns of wire carrying a current of 5 ampères. The permeability of the iron may be taken as 100.

8. Find the force in kilograms which an electromagnet can sustain when it is magnetized so that there are 600 lines of force per sq. cm. in the core, each pole piece having an area of 36 sq. cms.

Interaction of Currents and Magnets

695. Mutual Action of Parallel Currents. *Two parallel conductors carrying currents in the same direction attract each other, while if the currents are in opposite directions they repel.* This may be shown by means of Ampère's frame, a light rectangular frame of wire connected to a battery through two mercury cups so that it can freely revolve, as shown in figure 388. If a second frame having a number of turns of wire through which a current passes is brought up so that one of its edges is parallel and

near to one of the vertical wires of the pivoted frame, the attraction or repulsion of parallel currents is easily demonstrated.

Also if the frame B is held under the pivoted frame so that its upper edge is at right angles to the lower wire of the movable frame the latter will then turn until the two are parallel and with the adjacent currents in the same direction.
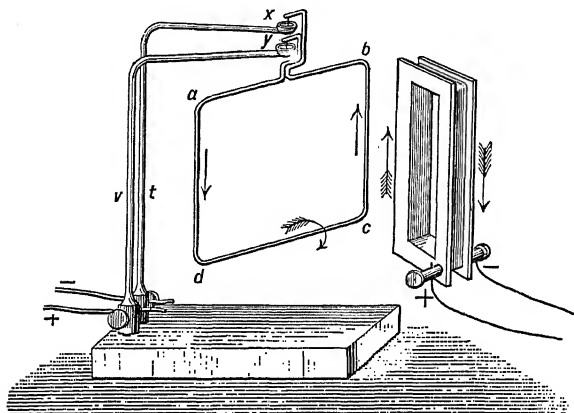
FIG. 388

696. Magnetic Field Around Parallel Currents. If the lines of force of two parallel currents are studied by means of iron filings or a compass needle, they will be found as in figure 389 when the currents are both in the same direction. While if the currents are in opposite directions the resultant lines of magnetic force are as shown in figure 390.

According to Faraday's conception, the attraction in the first case may be explained by a tension in the magnetized medium or a tendency for it to shorten up in the direction of lines of force; on the other hand, the repulsion in the second case is also in accordance with Faraday's idea that there is a pressure or tendency for a magnetized medium to expand at right angles to the lines of force.

It is also to be noticed that *in the first case the field of force is stronger just outside of the conductors than it is between them*, for between the two the magnetic effect of the one is opposed by the

other; while *in the second case the two act together to produce a strong magnetic field between them and a weaker field outside.*
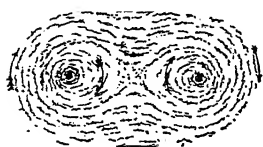
FIG. 389. Lines of magnetic force, currents perpendicular to paper and both down

FIG. 390. Magnetic field of two currents perpendicular to the paper; one down and one up

697. Action Between Current and Magnetic Field. In the case just considered each conductor may be thought of as acted on by the magnetic field due to the other. That there is such a reaction between a magnetic field and a conductor carrying a current may be demonstrated by presenting one pole of a bar magnet to one of the vertical branches of Ampère's frame (§ 695) when the wire will move across the lines of force of the magnet.

Or if a current of electricity is established in a light flexible conductor of tinsel cord hanging between the poles of a horseshoe
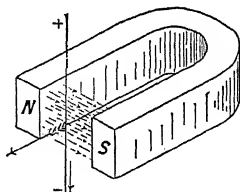
FIG. 391. Current in magnetic field
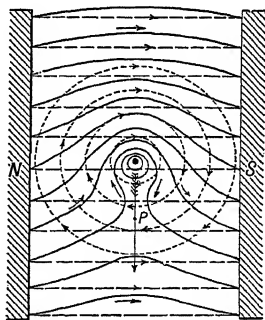
FIG. 392

magnet the cord is repelled outward from between the poles when the current is downward and the poles are situated as shown in figure 391. If the current is reversed or if the magnet is turned over so that the poles are interchanged the cord is drawn inward. The field of force due to a current flowing across a uniform magnetic field is shown in figure 392, where the current is supposed to

flow downward in a wire which intersects the paper perpendicularly at O. The broken lines are the lines of force of the uniform field, the circles are those of the current, and the full lines are the resultant lines of magnetic force. Clearly a tension in the medium along lines of force and pressure at right angles will urge the conductor in the direction shown by the arrow. At points nearer the top of the diagram than O, the force due to the currents acts *with* the original field, while below O, the two are in opposition, hence the field above O is strengthened by the current while it is weakened below the conductor, there being a neutral point P where there is no magnetic force at all.

These experiments lead to the following general rule:

When that magnetic field immediately adjoining a conductor carrying a current is strengthened on one side and weakened on the other by the effect of the current, the conductor is urged toward that side where the field is weakened.

If the current is not at right angles to the lines of force of the field, only *that component of the magnetic force which is perpendicular to the conductor* is effective, so that the effective magnetic force, the current, and the force acting on the conductor to move it are in three directions mutually at right angles to each other, and their relation can always be determined by the rule just given.

The amount of the force F experienced by the conductor is

$$F = lHI$$

where l is the length of the conductor in the field, H is the strength of the component of the magnetic field at right angles to the conductor, and I is the current strength, all being measured in C. G. S. electromagnetic units.

698. Magnet and Current in a Coil. The mutual action of a magnet and a current in a coil may be studied by a little light circular coil of wire connected to zinc and copper terminals which dip into a test-tube containing dilute sulphuric acid, the whole system being floated in a tank by means of a cork.* On present-

---

* In this experiment the test-tube should be weighted with shot until the cork is *entirely submerged*, only the upper part of the test-tube projecting above the surface, otherwise the motions will be greatly impeded by the surface viscosity of the water.

ing the pole of a bar magnet the coil will set itself so that the lines of force of the magnet are in the same direction through the coil as the lines of force of the coil itself, thus strengthening the field within the coil. It will then approach the pole and slip over it to the middle of the magnet.

If the magnet is now pulled out and quickly thrust through the coil in the opposite direction, the coil will slip off from the magnet, revolve so as to present the opposite face, and then again approach it.

It may be seen that these actions result from the general rule of § 697. For in each portion of the circular circuit the magnetic field is strengthened on one side of the conductor

FIG. 393. Magnet and floating current

and weakened on the other and each part strives to move from the stronger toward the weaker field. On the whole, therefore, *the coil always turns and moves so as to increase the resultant number of lines of force through it.* It slips to the middle of the magnet and then sets itself obliquely as shown in figure 394, for in that position it embraces the whole number of lines of force through the magnet and avoids also including those that turn back at the sides, which are in the opposite direction.

FIG. 394

FIG. 395. Barlow's wheel

699. Barlow's Wheel. In the apparatus shown in figure 395 a copper disc is balanced on an axle so that it can turn freely between the poles of a horseshoe magnet which produces

a strong field perpendicular to the disc. The lower edge of the disc dips in a trough of mercury. If one pole of a battery is connected to the axle of the disc and one to the mercury trough, a current will flow through the disc between its center and the trough. This current, being perpendicular to the lines of force of the magnet, is urged to the right or left, depending on its direction, and accordingly the disc itself is set in continuous rotation.

This experiment appears to show that the displacing force acts on the conductor which transmits the current and not simply on the current itself.

Instruments for Measuring Current and Potential

700. Measurement of Current. The strength of an electric current may be measured by its magnetic effect or by its heating or chemical action. Instruments which measure a current by its action on a magnetic needle are known as *galvanometers*.

701. Tangent Galvanometer. In the tangent galvanometer there is a circular coil having one or more turns of wire, at the
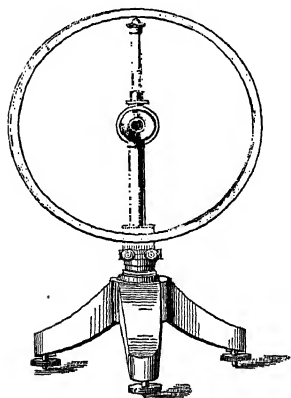
FIG. 396. Tangent galvanometer

center of which a magnetic needle is either balanced on a point or suspended by a fine fiber of silk or quartz. The instrument is placed so that the plane of the coil is vertical and in the magnetic north and south plane. When a current is sent through the coil the needle turns to one side or the other, and *the strength of the current is proportional to the tangent of the angle of deflection,* as may be shown as follows:

The force due to the current in the coil is at right angles to the plane of the coil at its center (§ 686) and the strength of the field at that point in a given coil is proportional to the strength of the current. Let *G* represent *the strength of field at the center due to the coil when unit current is flowing,* then *IG* will be the strength of field when the current

strength is I. Let OA in figure 397 represent the plane of the coil and O the point where the needle is placed, then when no current is flowing the needle points in the direction OA, being acted on only by the horizontal component H of the earth's magnetic force. The magnetic force F due to the current in the coil is IG and at right angles to H, therefore, the resultant force R is the diagonal of the rectangle whose sides are IG and H, and

$$tan\ x = \frac{IG}{H}$$
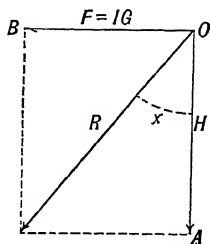
FIG. 397

where x is the angle which the resultant force makes with H. But the needle must point in the direction of the resultant force, and so x is the angle through which the needle turns. Therefore

$$I = \frac{H}{G}\ tan\ x$$

and if H and G are known the current may be determined by measuring the angle x.

702. Coil Constant of a Tangent Galvanometer. In case of a tangent galvanometer the magnetic force F due to the coil is expressed by IG.

But if the current is measured in electromagnetic units,

$$F = \frac{Il}{\cdot} \qquad G = \frac{\cdot}{r^2}. \qquad (\S\ 686)$$

And since the length of n turns of wire of radius r is $2\pi rn$,

$$\frac{2\pi nr}{} \qquad \frac{2\pi n}{r}.$$

The galvanometer *coil constant* G can be calculated from this formula when the coil of the galvanometer has so large a radius compared with the length of the needle that the poles of the needle may be regarded as at the center, and when the cross section of the coil is so small that all the turns bear nearly the same relation to the needle.

If G is determined in this way, r being measured in centimeters, and if H is found by the method described in § 510, the current will be found in

C. G. S. electromagnetic units by the use of the formula $I = \dfrac{H}{G}\ tan\ x$.

To obtain the current strength in *ampères,* we must take as the value of
the coil constant

$$G = \frac{2\pi n}{10.r}.$$

*By this method the strength of a current is determined in ampères directly
from the fundamental units of length, mass, and time,* for we have already
seen how the measurement of H is based on these units. A tangent galva-
nometer in which the constant is determined in this way directly from
measurements of the coil is known as a *standard galvanometer.*

703. Sensitive Astatic Galvanometer. For the measurement
or detection of extremely small currents of electricity the coil
of wire must contain a great number of turns as close as possible
to the needle, and because the turns nearest to
the needle are most effective it is customary to
use finer wire for these turns so that a greater
number can be placed in a given space.

The sensitiveness of the instrument is
further increased by using an *astatic needle.*
This is a system of two magnetic needles, as
nearly as possible of the same strength, con-
nected together by a light aluminum wire so
that the poles of the two needles are oppo-
sitely directed, as shown in the figure.

The combination is then suspended by a
fine silk or quartz fiber so that the galvanom-
eter coil surrounds only one needle, or the
second needle may be surrounded by another
coil around which the current flows in the
opposite direction to the first so that any
current in the coils will tend to turn both
needles in the same direction.

FIG. 398.   Astatic
   galvanometer
     diagram

If the two needles have equal magnetic moments the earth
will have no directive action on the combination. But as no
system of needles will remain perfectly astatic, a directive magnet
above or below the instrument serves to balance the effect of the
earth on the combined system. The influence of this magnet
and the torsion of the suspension fiber serves to give the needle
a definite position of rest.

A small mirror attached to the needle enables its angular

deflection to be measured by the usual telescope and scale method or by the reflection of light upon a scale.

704. Moving-coil Galvanometer. In this type of instrument, known also as the *D'Arsonval* form of galvanometer, the suspended system is a coil of fine wire which hangs in a strong magnetic field due to a permanent steel horseshoe magnet. In

figure 399 is shown a vertically placed horseshoe magnet, between the poles of which is hung a light rectangular coil of many turns of fine wire, the plane of the coil being parallel to the direction of the lines of force. The coil is suspended by a fine ribbon of phosphor-bronze which also serves to connect one end of the suspended coil to the outer circuit while the other connection is made through a spiral wound strip of the phosphor-bronze ribbon attached to the lower end of the coil.

A cylindrical mass of soft iron is fixed midway between the poles of the magnet so that as the suspended coil turns its vertical branches move in the gaps between the core and pole pieces. This arrangement secures a strong uniform field across which the

FIG. 399. D'Arsonval galvanometer

wires of the coil pass, and when a current is sent through it, it is deflected.

A small mirror mounted just above the coil and moving with it enables the deflection to be determined by the telescope and scale or reflected spot of light method.

The moving-coil galvanometer has the advantage that it is not affected by changes in the earth's magnetic field, and can be used near dynamo machines and where there is considerable magnetic disturbance. Also the coil damps strongly or comes almost immediately to rest when the wires leading to it are touched together, forming a *short circuit*, as it is called. This damping is due to electromagnetic induction (§ 730).

705. Electrodynamometers. Instruments in which no iron or

magnetic substance is used, but where the measurement depends on the mutual action of two coils carrying currents, are known as *electrodynamometers*.

The Siemens electrodynamometer, shown in figure 400, is a good example. An oblong coil is fixed in a vertical position and surrounding it closely at right angles, but not touching it, is a rectangular suspended coil. The current passes through the
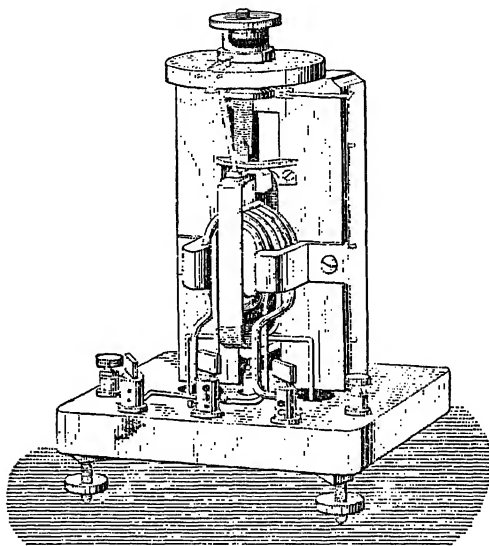
FIG. 400.　Siemens electrodynamometer

fixed coil and is led into the suspended coil through two mercury cups into which its ends dip. The magnetic action of the coils upon each other causes the suspended coil to turn, but its top is attached to a helical spring the upper end of which is fastened to a knob which is turned till the torsion of the spring forces the suspended coil back again into its zero position. The strength of the current is determined from the amount of torsion required, as shown by a circular scale.

In such an instrument the force of torsion T depends on the current strength in each coil or T is proportional to II', but when the current is the same in each coil T is proportional to I^2, whence

$$I = k\sqrt{T}$$

where k is a constant for the instrument which depends on the size, shape, and number of turns in the coils and the scale by which the torsion is measured. It is determined by experiment, by measuring the torsion produced by a known current.

When the current is reversed in both coils the deflection is in the same direction as before; for this reason *an electrodynamometer can be used to measure a rapidly alternating current which would give no deflection in a galvanometer.*

706. Ammeters. An *ammeter* is some form of galvanometer or electrodynamometer graduated so that the current strength in ampères may be directly read from the scale. A form of
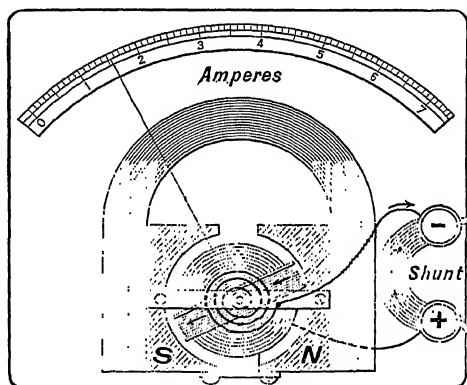
Fig. 401. Ammeter

ammeter much used for direct currents is shown in figure 401. It consists of a sensitive moving-coil galvanometer in which the coil instead of being suspended is mounted in jeweled bearings and is held in equilibrium by two non-magnetic spiral springs which also serve as conductors for the current. The main current passes through a strip of metal (called a shunt) having very small resistance, only a minute portion of the current passing through the delicate movable coil. But the current in the movable coil is always the same proportional part of the whole current, and therefore the scale over which the pointer moves

may be so graduated as to show directly the number of ampères in the *total* current.
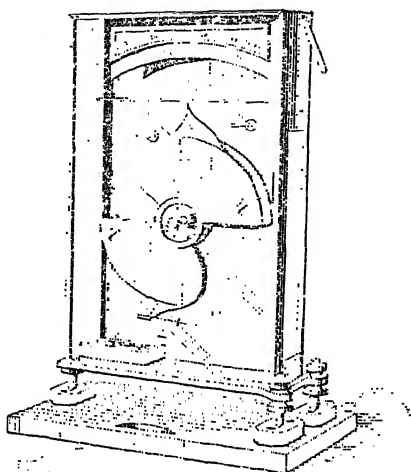
Fig. 402. Electrostatic voltmeter

An instrument of this type has the advantage of having a very small resistance.
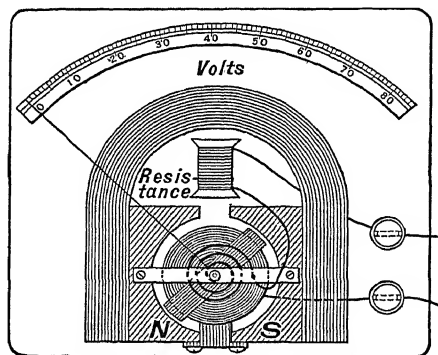
Fig. 403. Voltmeter

707. Voltmeters. A *voltmeter* is an instrument designed to measure differences in potential, and gives the readings directly

in volts. There are two principal types, *electrostatic* voltmeters and *those that depend on the flow of current.*

708. Electrostatic Voltmeters. These instruments are electrometers adapted to meet the requirements of ordinary engineering practice. Of this type is the instrument shown in figure 402.

709. Current Voltmeters. A voltmeter using current is a high-resistance galvanometer with a scale graduated to give directly the number of volts difference in potential between its terminals.

The voltmeter shown in figure 403 is a moving-coil galvanometer such as is used in the ammeter shown in figure 401, but there is no shunt across between the terminals as in the ammeter, and a considerable resistance is inserted in the circuit so that only a small current passes through the instrument.

Voltmeters using current give correct values only in circumstances where the current through the instrument is so small that it does not appreciably change the potentials to be measured.

For instance, the difference of potential of two statically charged bodies could not be determined by such an instrument, for they would be instantly discharged through it. And if we attempt to measure the difference of potential of the terminals of a battery cell whose internal resistance is as great as that of the voltmeter itself, the deflection will indicate only one-half the total electromotive force of the cell, for the current is such that half the fall in potential takes place in the cell itself (§ 650).

In ordinary commercial work the other resistances in the circuit are so small compared with that of a well-constructed voltmeter that there is no difficulty on this score.

Such a voltmeter cannot be used for alternating currents.

710. Ammeter with Iron Core. A simple form of ammeter is that shown in figure 404 in which a soft-iron core is drawn into a helical coil through which

Fig. 404. Ammeter with iron core

the current flows. Both ammeters and voltmeters are constructed on this principle, and as the soft-iron core is drawn

inward when the current is in either direction they may be used
either for direct or alternating currents, though the graduation
must be different in the two cases.

711. Hot-wire Instruments. In some instruments the cur-
rent passes through a fine wire and the elongation resulting from
its heating causes a pointer to move over a scale. The scale
may be graduated to show either the current in ampères or the
difference in potential between the terminals in volts. The wire
is mounted in a metal case to screen it from air currents and
keep it under as uniform conditions as possible.

The heating effect of a current is irrespective of its direction,
and therefore such an instrument may be used either for direct
or alternating currents.

712. Wattmeter. If it is desired to know the *energy per
second* or *watts* spent in any part of a circuit, as in the lamps
between A and B in the left diagram of figure 405, the current
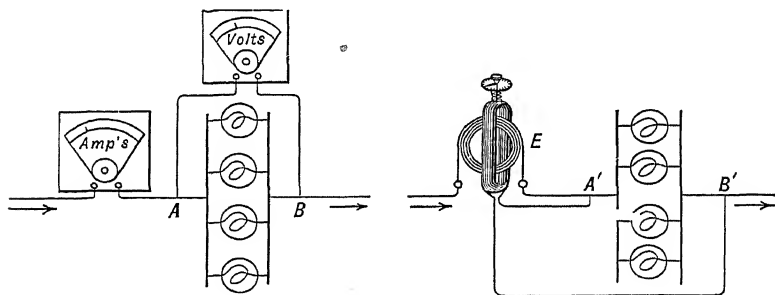
FIG. 405

may be measured by the ammeter and the difference of potential
between A and B by the voltmeter. The watts expended are
given by the product of the current in ampères by the volts.

The result may, however, be obtained directly by using a
wattmeter. This may be an instrument like the Siemens electro-
dynamometer connected so that the main current flows through
the fixed coil E (right diagram, figure 405) while the suspended
coil has a great many turns of fine wire and is connected at A'
and B' to the main circuit, so that the current in the suspended
coil will be proportional to the difference of potential in volts
between A' and B'. The torsion produced by the mutual action

of the two coils is proportional to the product of the currents in each, and is, therefore, proportional to IP where I is the main current and P is the potential difference between A' and B'. The instrument may therefore be graduated to give directly the watts expended between A' and B'. The suspended coil in this case is known as the potential or pressure coil, while the fixed one is the current coil.

PROBLEMS

1. What is the force of attraction between two straight parallel wires 30 cms. long and 1 cm. apart each carrying 3 ampères of current?

2. What must be the diameter of a coil of 3 turns of wire in order that a current of 5 ampères may produce a strength of field at its center of 0.20 dyne per unit pole?

3. A sensitive galvanometer having a resistance of 25 ohms is deflected one scale division by a current of $\frac{2}{1500}$ of an ampère. What resistance is required and how connected to change it into a voltmeter reading 1 volt per scale division, and what resistance and how connected to change it into an ammeter reading 1 ampère per scale division?

4. Given a voltmeter having a resistance of 800 ohms and reading 1 volt per scale division. How can it be made to read 10 volts per division?

5. How can the voltmeter described in problem 4 be used to find the current flowing through a conductor having a resistance 0.01 ohm per foot in length?

BELLS AND TELEGRAPH

713. Electric Bells. Bells are rung by electricity by the method shown in the figure. When the key at k is pressed, making a connected circuit, the current flows around the electromagnet M, causing it to attract the soft-iron armature a to which is attached the hammer which strikes the bell. But as the armature a is drawn toward the magnet a metallic contact at b is separated, thus interrupting the circuit and causing the magnet to lose its magnetism. The armature being mounted on a spring flies back, makes

FIG. 406. Electric bell

contact again at b, and is then again attracted by the magnet as at first.
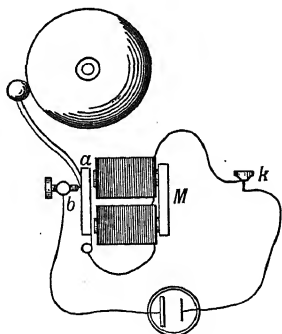
714. Electric Telegraph. In the Morse telegraph, as originally used, a recording instrument made a dot or dash when the key was pressed in the distant station. In this instrument a strip of paper was drawn steadily over a roller by clockwork, and when the key was pressed an electromagnet drew up a lever provided with a sharp steel point which pressed against the paper making a dot or dash, depending on whether the key made an instantaneous or more prolonged contact.

It was soon discovered that operators read the messages by sound, and therefore the elaborate recording instrument was replaced for the most part by the sounder, a simple electromagnet and armature arranged so that a vigorous *click* is heard when the circuit is closed or broken. In consequence of the resistance of long lines the current is very small and is therefore used to operate a *relay*, which merely closes the connection in a local battery circuit in which the sounder is included. The relay has a magnet wound with a great many turns of wire and in front of its poles is a nicely balanced armature controlled by a delicate spring so that a very small force will attract it. The armature is connected to the binding post a and the stop against which it is drawn is connected to b, so that when it is attracted by the influence of the main-line current, connection is made between a and b, thus closing the local circuit which includes the battery B and sounder S. The feeble motions of the relay armature are thus reproduced by the vigorous clicks of the sounder.

Fig. 407. Relay and sounder

Since the magnet of the relay must have a great many turns of wire, it must be wound with fine wire and will therefore have a large resistance; but since the resistance in the main line is already large, the additional resistance of the relay will have a comparatively slight effect on the current.

In the local circuit, the sounder and battery are all included in the same station and the resistance of the circuit may therefore be very small, hence the resistance of the sounder should be small, and accordingly it is wound with fewer turns of coarser wire.

In the main-line circuit a single wire of galvanized iron or hard drawn copper is used, the return circuit being through the earth. The following diagram shows the arrangement of a main line including three stations.

Each station has a key, relay, sounder, and local battery indicated, respectively, by k, R, S, L.

The main-line battery P_1 operates all the relays for a certain length of the line. At the last station shown in the diagram there is a relay R' which
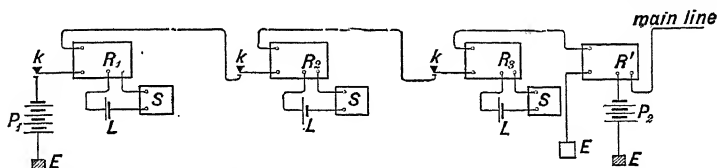
FIG. 408.   Diagram of telegraph line

transmits the signals to a second section of the main line which is operated by the battery P_2.

The keys are all provided with switches by which the circuit is kept closed everywhere except in the station where the operator is sending a message.

715. Cable Telegraphy. Ocean telegraphy presents some serious difficulties from which land lines are comparatively free. The cable acts like an enormous Leyden jar, for it consists of a central conducting core made of a bundle of copper wires twisted together surrounded by a thick coating of rubber insulating material, outside of which is a protecting sheath of hemp and steel wires.

The copper core is the inner coating of the jar and the steel sheathing is the outer coating. The capacity of an Atlantic cable is about equal to 600,000 gallon Leyden jars. When one end of the cable is connected to the battery the current at the other end rises to its full strength only very slowly, as the cable is being charged at the same time. And when the current is broken the whole charge has to escape before the current dies out. In a typical Atlantic cable the current rises to $\frac{1}{10}$ of its maximum value in 0.2 second, and would require 2 seconds to come to $\frac{9}{10}$ of its maximum; therefore, in order to save time, exceedingly sensitive instruments must be used which will give an indication as soon as the current begins to rise at the farther end. In giving a signal, connection is made to the battery for an instant and then the end is grounded, thus sending a sort of wave into the cable which is sufficient to affect the instrument at the other end without fully charging the cable.

A double transmitting key is used by which the cable may be connected either to the positive or negative pole of a battery, and thus a series of waves may be transmitted, positive corresponding to dots, and negative to dashes of the telegraphic code.

The receiving instrument is a sensitive galvanometer which swings to the right or left as the waves of current pass through it.

On a line connecting points so far apart on the earth there is a tendency for earth currents to flow which would powerfully affect the delicate galvanometers used and completely overpower the desired signals. To obviate this difficulty Varley devised the plan of connecting the cable at each end
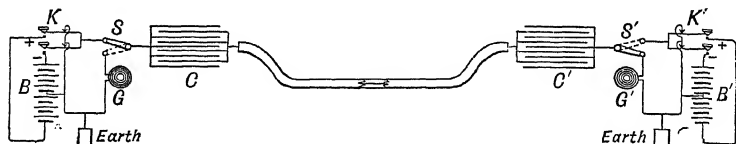
FIG. 409. Diagram of cable connections

to a condenser of large capacity which entirely prevents any steady flow through it due to earth potentials, but does not interfere with sending the signal waves.

A simple arrangement of a cable is shown in the above diagram.

The switches SS' are shown in position for sending by the key K and receiving by the galvanometer G'. Pressing the upper key at K gives a positive charge to the condenser C, while the other key gives it a negative charge. One terminal of the galvanometer G' is connected to the condenser C' while the other terminal is connected to earth.

716. Siphon Recorder. The instrument now commonly used for receiving cable messages is the *siphon recorder* devised for the purpose by Lord Kelvin. It is a galvanometer of the type which later became known as the D'Arsonval form. A coil of wire hangs between the poles of a powerful magnet, and through this coil the cable currents pass, causing it to turn. Attached to the suspended coil is a fine capillary tube of glass shaped like a siphon, one end of which dips into a little cup of ink. The other end of the siphon tube just touches a strip of paper which is carried along by clockwork. As the coil turns the siphon moves to and fro across the paper, tracing a wavy line as the paper moves along. An automatic jarring apparatus prevents the friction between the paper and point of the siphon from interfering with the free motion of the coil.

ELECTROMAGNETIC INDUCTION

717. Faraday's Discovery. The year 1831 was made memorable by the discovery of electromagnetic induction by Michael Faraday, then professor in the Royal Institution in London. In seeking to find some action of an electric current on a neighboring conductor Faraday, having placed a coil of wire carrying an electric current upon another coil which was connected to a galvanometer, found that if the electric current was interrupted or broken there was a sudden deflection of the galvanometer

lasting only for an instant, and when the battery connection was made again there was an equal deflection but in the opposite direction. But the steady flow of current in one coil had no effect whatever upon the other.

These momentary currents are called *induced* or *secondary* currents, while the battery current by which they are produced is called the *primary* current. The corresponding coils of wire are known as the primary and secondary coils.

718. Induction by a Moving or Varying Current. Faraday also showed that when a coil carrying a current is moved either toward or away from another coil connected to a galvanometer, an induced current is set up.
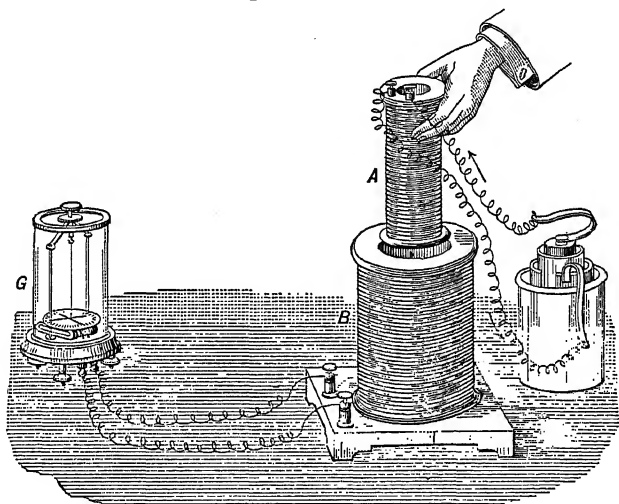
FIG. 410. Induction by a moving current

Such an arrangement as shown in figure 410 may be used, where the primary coil A has a current flowing through it from the battery and the secondary coil B is joined to the galvanometer. If the coil A is either pushed down inside of the coil B or withdrawn from it, an induced current is obtained which flows around B in the opposite direction to the current in A when the two are pushed together, but in the same direction as in A when the coils are drawn apart.

If while the coil *A* is inside coil *B* the current in *A* is made weaker, an induced current is set up the same as though *A* were being withdrawn. But when the current in *A* is strengthened the effect is as though the coils were moved closer together.

719. Induction by Magnets. Since a coil of wire carrying a current is surrounded by a magnetic field, it may be supposed that a magnet will produce a similar effect, and experiment shows this to be the case. When a bar magnet is thrust into a coil of wire connected in circuit with a galvanometer there is an instantaneous swing of the needle of the galvanometer, but *the needle at once returns to its zero position and remains there as long as the magnet is held at rest;* when it is withdrawn from the coil there is another instantaneous deflection opposite to the first. If the experiment is repeated with the magnet reversed, the deflections are opposite to those previously obtained.
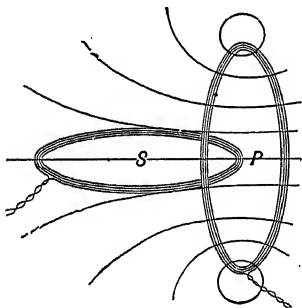
FIG. 411.   Coils with no mutual
                  induction
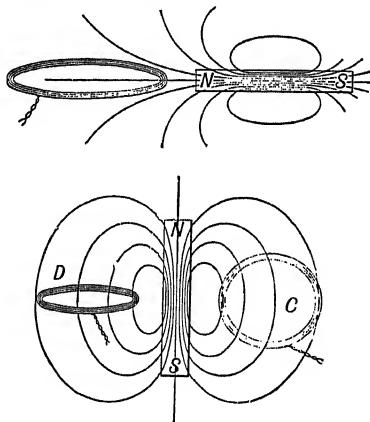
FIG. 412

720. General Condition of Induction. In general *an induced current is set up in a coil whenever there is a change in the number of lines of magnetic force passing through the coil.* This condition is illustrated in each of the three modes of producing induced currents just described. When the two coils of Faraday's first experiment are placed in the relation shown in figure 411 so that the lines of force due to the primary coil *P* instead of passing through the secondary coil pass on each side of it,

there is no induced current in the secondary coil. So also there is no induction when a magnet is brought up to the coil in the position shown in the upper diagram of figure 412 or when the plane of the coil is parallel to the magnet as shown in the coil C on the right of the magnet in the lower diagram, but when the coil is at right angles to the magnet as in the left-hand coil D there will be an induced current when the magnet is brought up or taken away, because more lines of force of the magnet pass downward through the coil when it is near the magnet than when it is at a distance.   (See § 511 on number of lines of force.)

721. Induction by Earth's Field.   The inductive effect of the earth's magnetism may be easily observed by means of a coil of large area and many turns of wire connected with a suitable galvanometer.

If such a coil is held with its plane perpendicular to the lines of the earth's magnetic force as at A, figure 413, the maximum number of lines of force will pass through it.   If it is now turned quickly into the position B parallel to the lines of force, where none pass through it, there is an induced current because of the change in the number of lines of force through the coil.   If the coil, instead of being turned half-way, is turned completely over, its position relative to the lines of force is exactly reversed and the inductive effect is twice as great as when it was turned half-way over.
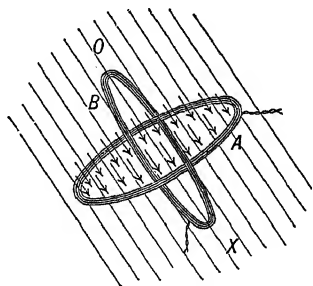
Fig. 413.   Coil in earth field

When the coil in any position is rotated about an axis OX parallel to the lines of force of the field, there is no induction since no change takes place in the number of lines of force passing through it.

When the coil is laid flat on a table and slipped about from one place to another there is no induction, even if the table is tipped so that its top is at right angles to the lines of force, because the same number of lines of force pass through the coil wherever it is, since the field is uniform.

722. Faraday's Disc.   The following experiment due to Fara-

day shows that *when a conductor moves across the lines of force of a magnetic field an induced electromotive force is developed.*

A copper disc is mounted on an axis so that it can rotate between the poles of a horseshoe magnet, the axis of the disc being parallel to the lines of force.   The edge of the disc dips into a

mercury trough connected to one end of a low-resistance galvanometer circuit, the other end of which is put in contact with the axle of the disc.
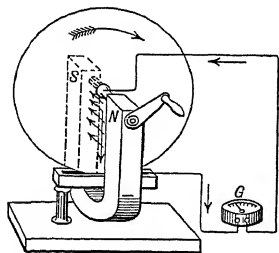
On rotating the disc in the direction of the arrow a current is set up in the direction shown in the figure, the strength of which is proportional to the speed of revolution of the disc. If the disc is rotated in the opposite direction the current is reversed.

FIG. 414.   Faraday's disc

This experiment shows that *each radial strip of the disc, as it cuts across the lines of force of the magnetic field, is the seat of an electromotive force which is found to be proportional to the number of lines of force cut across per second,* for it is proportional both to the speed of rotation of the disc and to the strength of the magnetic field.

723. Electromotive Force of Induction.   When the C. G. S. electromagnetic system of units is used (§§ 616–617) *the electromotive force of induction is numerically equal to the number of lines of force, or unit tubes, cut across per second by the conductor;* that is,

$$E = \frac{N}{t} \qquad (E \text{ in } C. G. S. \text{ electromagnetic units})$$

where E is the electromotive force induced in a conductor which is cutting across lines of force *at the rate* of N lines in t seconds.

Or, since one volt (§ 617) is equal to 10^8 C. G. S. electromagnetic units of potential,

$$E = \frac{N}{10^8 t}. \qquad (E \text{ in volts})$$

To prove this relation suppose a circuit, such as is shown in figure 415, consisting of two straight parallel conducting rails connected together at one end and also connected by a cross conductor AB which can slide in the direction of the arrow; and let this circuit be in a magnetic field of strength

H in which the lines of force are perpendicular to the plane of the circuit. Then if AB is slid along by hand at the rate of x cm. per second, an induced electromotive force will be produced which will cause a current I in the circuit.

FIG. 415

The energy expended per second by this current will be IE (§ 664), but this energy is supplied by the work expended in moving the conductor along and must be equal to it. But a conductor of length l which carries a current I across a magnetic field of strength H is acted on by a force $F = HIl$ (§ 697) and if in one second the conductor is moved against that force through a distance x the work done in one second is $Fx = HIlx$. We have then,
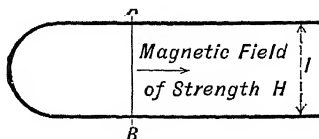
$$IE = HIlx \quad \text{or,} \quad E = Hlx,$$

but lx is the area moved over by the conductor AB in one second, and so Hlx equals the number of lines of force cut across per second.

Thus the electromotive force of induction in C. G. S. electromagnetic units is shown to be numerically equal to the number of lines of force cut across per second by the moving conductor.

724. Illustration. For example, suppose a straight conductor AB, one meter long (Fig. 416), is moved in the direction of the large arrow at the rate of 3 meters per sec., and suppose it is in a magnetic field of strength 0.5 (about as strong as the earth's field) in which the lines of force are *down* perpendicular to the paper. Then the number of lines of force cut per second will be the area in centimeters swept across per second by the conductor, multiplied by the number of lines of force per square centimeter which in this case is 0.5, or $E = 100 \times 300 \times 0.5 = 15000$, which is the electromotive force in C. G. S. electromagnetic units; to change it to volts it must be divided by 10^8, hence
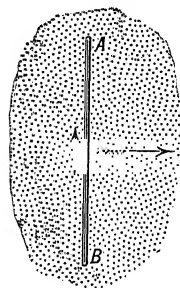
FIG. 416. Wire moving across lines of force

$$E = 0.00015 \text{ volt}$$

which is *the difference of potential between the ends of the wire,* since it is disconnected and no current can flow.

725. Why Induction Depends on Change in Number of Lines of Force through a Circuit.

We are now prepared to understand why it is that the resultant electromotive force induced in a circuit depends on the *change* in the number of magnetic lines of force passing through the circuit.

It has already been seen that when a coil of wire lying on a

table is slid along, no induced current is produced although the wires of the coil cut across the lines of force of the earth's magnetic field (§ 721). The explanation of this is that *electromotive forces are induced, but in such a way that they balance each other.*
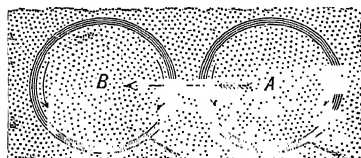
FIG. 417. Coil moved sidewise in a magnetic field

For suppose the coil is moved from A to B as in figure 417 and that the lines of magnetic force are straight down perpendicular to the diagram, then the sides of the coil cut across lines of force in such a way as to cause electromotive forces in the direction of the arrows. The electromotive forces induced in the two sides therefore act against each other in the ring, but they are equal because each side of the ring cuts across the same number of lines of force in the same time, therefore the electromotive forces balance and there is no current.

If the field is not uniform so that more lines of force pass through the coil in the second position B than in the first position, then more lines of force must have cut into the coil across its left-hand side than have cut out of it across its right-hand side. The electromotive force developed in the left-hand side of the coil will then be greater than the other and will cause a current to flow around the coil counter-clockwise. Therefore there must be a resultant electromotive force whenever the number of lines of force through a coil is increased or diminished.

726. Induced Electromotive Force. Since the electromotive force developed in any part of a conductor by induction is equal to the number of lines of force which cut across it per second (§ 723), it follows that *in any circuit or coil the electromotive force of induction is equal to the change per second in the number of lines of force included by the circuit.*

This is expressed by the formula

$$E = \frac{N_1 - N_2}{t}$$

which gives the *average* electromotive force during the time interval t when N_1 is the number of lines of force through the

circuit at the beginning of the interval and N_2 the number at the end.

By taking the time interval very short we approach the instantaneous value of the electromotive force as a limit.

If there are several turns of wire in the coil, to get the total electromotive force the above expression must be multiplied by the number of turns.

It is clear from the above that the more quickly the change in the number of lines of force takes place the greater the electromotive force.

727. Induced Current and Total Flow. The induced current at any instant is by Ohm's law

$$I = \frac{E}{R} \quad \text{and since} \quad E = \frac{N_1 - N_2}{t}$$

we have
$$I = \frac{N_1 - N_2}{Rt}$$

The instantaneous value of the induced current is therefore greatest when the induced electromotive force is greatest; that is, when the change in the number of lines of force through the circuit is taking place most rapidly.

But It, the product of current by the time that it flows, is the whole quantity of charge or electricity that passes in time t; thus

$$It \text{ or } Q = \frac{N_1 - N_2}{R}.$$

A simple integration shows that this expression holds true in every case, at whatever rate the lines of force through the circuit may be changing.. *The total quantity of electricity passing a given point in the circuit in consequence of induction is equal to the change in the number of lines of force through the circuit divided by its resistance.* If C. G. S. units are used for N and R the quantity Q will also be in that system. To find it in coulombs it must then be multiplied by 10.

It is to be remarked that *the total quantity of the induce flow is independent of the time during which the induction takes place.* It is the same when a magnet is put into a coil as when it is pulled out and whether it is moved slowly or rapidly.

728. Energy in Induction. Every current of electricity possesses energy, and therefore energy is required to produce induced currents. During the changes which produce an induced current energy is supplied to it,, and it dies out immediately when the inductive action stops because its energy is expended in heat in the conductor if in no other way. When induced currents are set up by making or breaking the current in an adjoining primary circuit the energy comes from the primary battery. When the induced current is caused by the motion of a conductor in a magnetic field the energy is supplied by the agency which causes the motion.

For instance, more energy must be expended when a magnet is thrust into a coil in which the ends of the wire are connected forming a closed circuit than if the ends had not been joined, for there is an induced current in the first case and not in the other. But in order to expend energy resistance must be overcome, and so the induced current must cause a force which resists the magnet as it is pushed into the coil. For the same reason the current which is induced when the magnet is withdrawn must exert a force to resist the withdrawal of the magnet so that more work is done than if the current could not flow.

729. Lenz' Law. The general law suggested in the last paragraph was first stated by Lenz and is known by his name. It may be stated thus: *An induced current is always in such a direction as to resist by its electromagnetic action the motion by which it is produced.* This law is a direct consequence of the conservation of energy, as has been already indicated.

730. Illustrations of Lenz' Law. Thus in case of Faraday's disc experiment (§ 722) the induced current tends to rotate the disc in the opposite direction (see Barlow's wheel, § 699) so that it is harder to turn the disc while the induced current is flowing than if the circuit were disconnected.

If a thick strip of sheet copper is hung like a pendulum so that it can swing edgewise between the poles of a powerful electromagnet, it may swing down with a rush, but is instantly checked as it comes between the magnet poles, since there are induced in the copper currents of electricity which resist the motion, transforming the energy of motion into current energy which finally results in heat in the copper. In some forms of galva-

nometer a bell magnet is employed, so called because it is shaped like a cylindrical bell of steel slit part way up, the poles being on the two sides. If such a magnet is suspended in a slightly larger cylindrical cavity in a copper block, it generates by its motion induced currents which quickly bring it to rest. This mode of stopping the vibrations of a magnetic needle is called *electrical damping*. The damping of the coil of a D'Arsonval galvanometer (§ 704) is also explained in the same way.

731. Arago's Disc. A celebrated experiment of Arago's, which was first explained by Faraday, is illustrated in figure 419. A copper disc is rotated rapidly under a magnetic needle from which it is separated by a sheet of glass or parchment which prevents air currents from having any influence on the needle,

FIG. 418.   Copper pendulum and magnet

FIG. 419. Currents in Arago's disc

and the needle is carried around with the disc. Induced currents are set up in the disc which resist the relative motion of the two, consequently the needle is dragged along after the disc. The lines of force due to the needle go down through the disc under the north pole and the induced currents are as indicated by the dotted curves. It is easily seen that the current flowing under the needle will tend to cause it to turn in the direction of the disc, as in Oersted's experiment (§ 683).

732. Rules for the Direction of Induced E.M.F. and Current. Lenz' law leads to the following rules for the direction of the

induced electromotive force and resulting current:

Case of a Wire Moving Across Lines of Force

In this case *the electromotive force induced in the wire is in such a direction as to cause a current which will strengthen the field immediately in front of the moving wire and weaken the field immediately behind it.*

For it has been seen in § 697 that such a current would urge the wire across the field in the opposite direction, thus resisting the motion.

Case of a Closed Circuit

When the number of lines of force through a circuit is increasing, the induced current is in such a direction as to set up lines of force through the circuit opposite to those already there, thus opposing the increase.

If the number of lines of force is decreasing, the induced current is in such a direction as to set up lines of force inside the coil in the same direction as those already there, thus opposing the decrease.

733. Self-induction or Inductance. When a current of electricity is set up in a coil of wire each turn in the coil experiences the inductive effect of the current starting in all the other turns. All act together to cause *an induced current in the coil opposite to the current which is starting.* The resultant current is therefore weaker than the steady current which will flow when the inductive action is over. When the circuit is broken *the self-induced current is in the same direction as the current which has been flowing;* it acts therefore *with* that current and prolongs
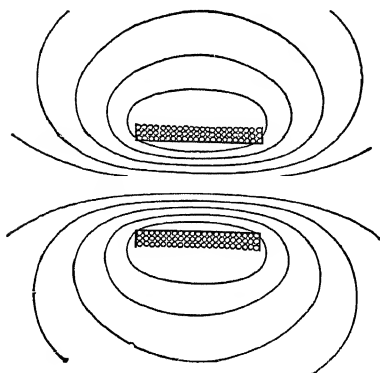
FIG. 420

its flow, causing a bright spark across the gap where the circuit is broken. The current induced on breaking connection is known as *the extra current.*

In this case, as in all other cases of induction, the action is due to that relative motion of conductors and magnetic field expressed by the phrase "cutting lines of force." The coil after the current is established has a magnetic field, and includes a large number of lines of force. These lines of force form closed curves surrounding the coil and may be considered as starting in the coil and spreading out in expanding curves as the current becomes stronger. Each turn of wire in the coil is cut by all the lines of force and hence *the electromotive force of self-induction depends on the number of turns of wire in the coil and the total number of lines of force that are set up by its current.* What is called the *coefficient of self-induction* or the *inductance* of a coil is the product of the number of its turns of wire by the number of lines of force through the coil when unit current is flowing in it. Thus even a circuit consisting of a single turn of wire has some self-induction, but it is greatest in coils which have many turns of wire and include a great number of lines of force, as in electromagnets, where the iron core immensely increases the self-induction.

734. Experimental Illustration. Take a large electromagnet of low resistance having an armature across its poles and connect a small incandescent lamp across its terminals, as shown in figure 421. Then join the magnet to a storage battery which is strong enough to light up the lamp when connected to it alone, interposing a contact key. On pressing the key the lamp lights for an instant as the electromotive force of self-induction opposes the flow of current through the magnet and sends it through the lamp instead. The lamp dies out, however, as the current comes to its steady state and divides between

FIG. 421. Self-induction

lamp and magnet. On breaking the circuit the lamp again glows as the self-induced current rushes around through the lamp instead of leaping the gap at the key. *The phenomena of self-induction are observed only while the current is changing,*
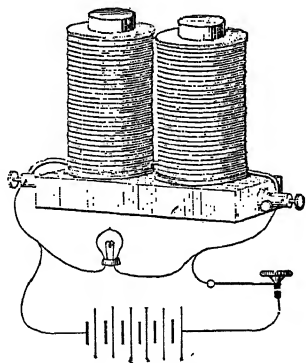
hence in case of steady currents self-induction need not be considered, but in dynamo machines and all alternating current apparatus it plays a most important part.

In breaking connection in a circuit containing much self-induction, such as one including electromagnets or a dynamo machine, great care must be taken not to be touching the conductors on both sides of the gap when the contact is broken; otherwise a severe shock may be obtained from the *extra current* even when the ordinary voltage in the circuit is small.

735. Energy of a Magnet. *Every portion of the magnetic field whether within the iron core of the magnet or outside of it has a certain energy in consequence of its magnetization.* It was shown by Maxwell that the energy per cubic centimeter in any part of a magnetic field is $\dfrac{B^2}{8\pi\mu}$, where B is the induction at that point or number of lines of force per square centimeter.

Therefore, when a current is starting in a coil or electromagnet it has to supply the energy of the magnetic field besides spending energy in heat owing to the resistance of the conductor. After the magnetic field is fully established, which may take several seconds in a large magnet, the current is steady and spends energy only in heat in the conductor. *No energy is required to keep up a magnetic field when it is once established.*

The spending of energy by a current in making a magnetic field causes the current to delay in coming to its full strength and is the cause of the self-induced current on making connection.

When the circuit is broken the field loses its magnetization and therefore gives up its energy again to the current. This causes the *extra current* or induced current on breaking the connection, and *the energy of this extra current is equal to the energy that was stored up in the magnet and surrounding magnetic field.*

736. Induction Coil. Ruhmkorff Coil. The induction coil is a device for obtaining induced currents of very great electromotive force from an ordinary battery current. The construction is illustrated in figure 422. The *primary coil,* of a few layers of large copper wire so as to have small resistance, is wound about a central core which consists of a bundle of soft-iron wires. Outside of the primary coil and thoroughly insulated from it by a thick tube of hard rubber is the *secondary coil,* made of an im-

mense number of turns of fine wire the ends of which are brought to two insulated posts supporting the discharging rods *a* and *b*. The diagram, for distinctness, shows only a few turns of wire in the secondary; but in the actual instrument there are thousands of turns, a coil to give a one inch spark must have something like a mile of wire in its secondary coil. The primary must be thoroughly insulated from the secondary by a thick tube of hard rubber with hard rubber flanges at the ends. The primary coil is connected to a battery of a few storage cells and when the current is interrupted the induced electromotive force
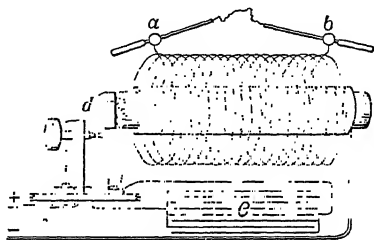
FIG. 422. Induction coil

in the secondary coil may be great enough to cause a discharge across between the discharging rods. *The primary current is automatically connected and broken.* A device commonly used is shown at *d*. A little block of iron on the end of a spring is mounted opposite the end of the iron core of the apparatus. The spring rests against the end of an adjusting screw, the points of contact on each of them being made of platinum. The connections are made so that the primary current flows across the contact between spring and screw, and consequently, as the core becomes magnetized and attracts the block of iron mounted on the spring, the connection is broken. But as the core then loses its magnetism the spring comes back and again makes the connection; and so the action is repeated, automatically making and breaking the current many times in a second.

The self-induction of the primary coil causes both the starting and stopping of the current to be prolonged, and consequently the E.M.F. of induction would be comparatively small if this were not obviated. It is found that if a condenser of suitable capacity is connected to the primary circuit, its two surfaces being connected one on each side of the point where the current is broken, the electromotive force produced on breaking is greatly increased. Such a condenser is represented at *C;* it is usually made of alternate sheets of tinfoil and paraffined paper, the odd sheets of tinfoil being connected together for one coating and the

even sheets forming the other. By this construction a large capacity is obtained in very compact form. The condenser is often contained in the base of the instrument.

When the current is broken at d the extra current of self-induction rushes into the condenser and charges it instead of discharging in a spark across the gap at d. The flow of the extra current is thus very quickly stopped; but after the condenser is charged it immediately discharges itself back through the coil in a direction opposite to the original current, and so more perfectly demagnetizes the core or even magnetizes it oppositely.

By the use of the condenser, then, there is a greater change in the number of lines of force on breaking the current, and the change is more instantaneous, both effects serving to increase the electromotive force of induction; and at the same time the sparking at the gap d, which is very destructive to the platinum contacts, is greatly reduced.

737. Wehnelt Interrupter. Instead of a mechanical interrupter for the circuit an electrolytic cell may be used, known as the Wehnelt interrupter from its discoverer. This cell consists

Fig. 423. Electrolytic interrupter

of a vessel containing dilute sulphuric acid, having for the negative electrode a plate of sheet lead and for the positive electrode a platinum wire or rod covered with a glass or porcelain sheathing so that only the tip end projects into the acid. An adjusting screw is provided by which the amount projecting may be regulated. If the voltage in the circuit is sufficient and the exposed tip of platinum wire is properly proportioned to the current, the circuit will be rapidly interrupted, bubbles of gas being given off at the platinum wire accompanied by flashes of light.

The frequency of interruption depends on the self-induction of the circuit and the electromotive force of the battery as well as upon the adjustment of the platinum point, and may be varied through wide limits. With this form of interrupter a condenser is of no advantage.

738. Telephone. In the early telephone as devised by Bell the receiver and transmitter were alike, the construction being shown in figure 424. A hard rubber handle contains a hard-steel cylindrical magnet, around one end of which is fixed a coil of many turns of fine wire the ends of which are brought to binding screws on the handle. A disc of thin sheet iron, supported at the edges so that it is free to vibrate in the middle, is mounted so that its center comes
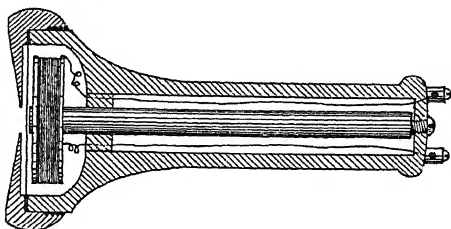
Fig. 424. Telephone receiver

close to the end of the magnet and surrounding coil but does not touch them. A hard rubber cap or ear-piece having a hole in the center fits over the disc and serves to clamp it firmly at the edges as well as to improve the quality of the tone by favoring the sound waves from the center of the disc.

Suppose two such instruments with the coil in one connected in closed circuit with the coil in the other. If a person speaks into one the sound waves impinging on the center of the disc cause it to vibrate; but as it vibrates induced currents are set up, for when the disc approaches the magnet more lines of force pass through the coil into the disc, and as it springs away the lines of force spread out again cutting across the coil. These induced currents flow through the coil around the magnet of the receiving telephone and by alternately opposing and strengthening its magnetism cause the iron diaphragm of the receiver to vibrate in exact correspondence with that of the transmitter, so that the same motion is given to the air at one end as that which caused the disc to vibrate at the other, thus reproducing the sound.

There is a serious defect in this mode of transmission. All the energy of the induced currents must come from the sound waves which cause the disc of the transmitter to vibrate, and as a part of this energy is spent in heat in consequence of the resistance of the circuit, the sound heard at the receiver must be faint. To meet this difficulty another form of transmitter

shown in figure 425 is ordinarily used.   The cell C containing carbon granules between two plates of polished carbon is mounted
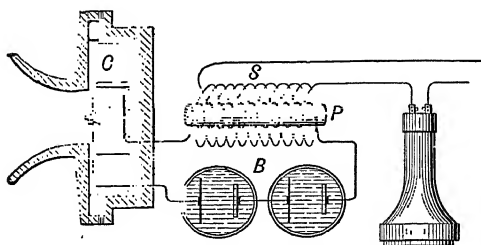
FIG. 425.   Transmitter

between the thin metal diaphragm and the solid back of the instrument, and on each side of the cell is a metal plate connected in circuit with a battery B and the primary winding P of a small induction coil, of which the secondary S is connected to the line leading to the receiving station.

When sound waves fall upon the diaphragm of the transmitter the vibrations cause a variation in its pressure on the carbon cell and a consequent change in its resistance.   The other resistances in the battery circuit are small compared with that of the granular carbon, hence variations in its resistance cause decided changes in the strength of the current.   These set up induced currents in the secondary and produce corresponding vibrations in the diaphragm of the receiver, thus reproducing the sound.
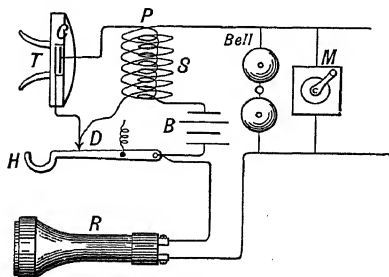
By this arrangement the energy for transmission is supplied by the battery, and by taking a proper number of turns in the secondary coil the induced current can be adapted to the resistance of the line.

FIG. 426

A telephone line is usually a complete circuit of two wires instead of using the earth, as in telegraphy, and the two wires are carried near together so that the inductive action of neighboring telegraph lines and lighting wires on one wire may be neutralized by their action on the other.

The arrangement adopted in the local battery system is shown in figure 426. When the receiver R is hung on the hook H the battery circuit is broken at D and also the secondary circuit so that no current flows from the battery except when the line is in use. The call bell is of very high resistance so that only a very small part of the current is diverted through it, and the magneto M by which the bell is rung is so devised that it is connected to the line only while being used.

The local battery is often done away with and the current through the subscriber's transmitter supplied by a single battery at the central station. One method of connection is shown in figure 427.

A battery B of about 24 volts is connected to the line at the central station; but when the receiver R hangs on the hook h there is no current in the line, for the circuit is broken at q and no current can flow across through the call bell M because the condenser e is interposed. The sub-
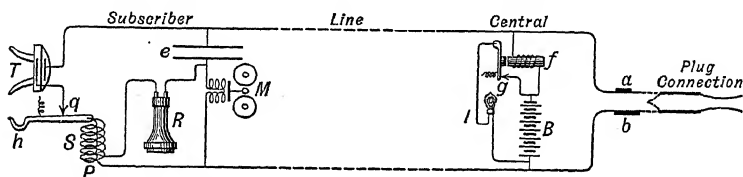
Fig. 427. Telephone with central battery

scriber may be called, however, by connecting to the terminals ab any source of *alternating* current, which causes a surging of current back and forth in the line to the condenser, charging it so that first one of its coatings is positive and then the other, alternately. This current as it flows alternately into the condenser and out again rings the call bell M. On the other hand, if the subscriber wishes to call "central" he has only to lift his receiver from the hook. The current is then established through the contact points at q and flowing around the relay f closes at g the circuit through the signal lamp l which flashes out and shows that a connection is desired. The correspondent's line is then connected by means of a flexible cord having two conductors and terminating in a double contact plug which connects one conductor to a and the other to b.

The receiver R is so connected that only an extremely minute *direct* current from the battery can flow through it on account of the large resistance of the bell M (1000 ohms), but the *alternating* "talking" current induced in the secondary s is readily established through the condenser e. The "talking" current and the direct current from the battery through the transmitter are thus both transmitted over the same line without interfering with each other.

ELECTROMAGNETIC UNITS

739. C. G. S. Electromagnetic Units. The C. G. S. electromagnetic system is based on the unit magnet pole as defined in

§ 497, unit current as in § 616, and unit electromotive force as in § 617. These units are determined from the above definitions by certain measurements of length in centimeters, of mass in grams, and of time in seconds. They have the advantage of being directly connected with the fundamental mechanical units of the C. G. S. system. Thus the product of current by electromotive force measured in these units gives the rate of spending energy in ergs per second.

The C. G. S. electromagnetic unit of resistance is the resistance of a circuit in which the above unit electromotive force will produce unit current of the same system.

740. Practical System of Units. The C. G. S. electromagnetic units are not of a convenient size for use in commercial measurements, but it is desirable that the practical units should be related to the C. G. S. electromagnetic units by ratios which can be expressed by simple powers of 10.

Thus *the volt, the practical unit of electromotive force, is chosen equal to* 10^8 *C. G. S. electromagnetic units of electromotive force,* because that particular power of 10 gives a value nearer to the electromotive forces of ordinary battery cells than any other would have done.

The ohm is defined as 10^9 *times the C. G. S. electromagnetic unit of resistance.* This power of 10 was adopted because it corresponds very nearly to the Siemens unit* of resistance which was already in use and had been found convenient.

The ampère is 10^{-1} *times the C. G. S. electromagnetic unit of current.* It is determined by Ohm's law as the current which results from an electromotive force of 1 volt in a circuit having a resistance of 1 ohm.

The coulomb is the unit of charge. It is the charge transmitted in 1 second by a current of 1 ampère. It is almost exactly equal to three thousand million electrostatic units of charge as defined in § 540.

The farad is the unit of capacity. It is the capacity of a condenser which will hold a charge of one coulomb when the differ-

---

* The Siemens unit is the resistance of a column of pure mercury 1 meter long and 1 sq. mm. in cross section, at the temperature of melting ice. Named from Sir William Siemens, the distinguished German physicist and engineer who advocated it.

ence of potential between its coatings is 1 volt. This unit is so large that ordinary condensers are rated in microfarads, or millionths of a farad.

*The henry is the unit of inductance.** It is the inductance of a circuit in which an increase in current strength at the rate of 1 ampère per second produces a back electromotive force of 1 volt.

Elaborate experiments have been made to determine how the units as above defined may be realized in practice, and the following experimental values have been obtained:

The ohm is the resistance of a column of pure mercury 106.3 cms. long and 1 sq. mm. in cross section at the temperature of melting ice.

The ampère is a current which will deposit 0.001118 gm. of silver per second in a silver voltameter.

The volt may be determined from a Weston normal cell, the electromotive force of which at 20° C. is found to be 1.0183 volts.

741. To Change from the Electrostatic to the Electromagnetic System. The ratio of any electrostatic unit to the corresponding electromagnetic unit is in every case some power of the velocity of light (3×10^{10}) cm. per second.

Electrostatic quantity of charge $\div (3 \times 10^{10})$ = charge in C. G. S. electromagnetic units.

Electrostatic quantity of charge $\div (3 \times 10^9)$ = coulombs.

Electrostatic potential $\times (3 \times 10^{10})$ = potential in C. G. S. electromagnetic units.

Electrostatic potential $\times 300$ = volts.

Electrostatic capacity $\div (3 \times 10^{10})^2$ = capacity in C. G. S. electromagnetic units.

Electrostatic capacity $\div (9 \times 10^{11})$ = farads.

Electrostatic capacity $\div (9 \times 10^5)$ = microfarads.

PROBLEMS

1. A coil of wire of 10 turns, each turn enclosing an area of 900 sq. cms., is turned from position *A* to *B* (see Fig. 413) in ½ second. Find the induced E.M.F. in volts when the strength of the magnetic field is 0.5.

2. A metal spoke in a wheel is 80 cms. long. If the wheel makes 300 revolutions per minute in a plane perpendicular to the lines of force of the

---

* Named in honor of Joseph Henry, a distinguished American physicist and first Secretary to the Smithsonian Institution, who discovered self-induced currents.

earth where the field strength is 0.5, find the difference of potential between the center and rim of the wheel. Which part is at the higher potential when the wheel rotates clockwise as seen by one looking in the positive direction of the lines of force?

3. Show that the work expended in producing an induced current by turning a coil over in a magnetic field becomes twice as great when the time of the operation is reduced $\frac{1}{2}$.

4. A railway train runs south on a straight track with a velocity of 25 meters per sec. If the vertical component of the earth's magnetic force is 0.50, find the electromotive force induced in a car axle 120 cms. long; also which end, east or west, is at the higher potential.

5. When the vertical component of the earth's magnetic force is 0.50, find the electromotive force induced in a coil of 10 turns of wire 3 meters in diameter which while lying on the ground is in $\frac{1}{2}$ second pulled out into a loop so long that the sides touch.

6. When a circular coil of 100 turns of wire 1 meter in diameter lying on the floor is turned over in 0.3 second, find the average electromotive force, earth field being as above.

7. If the resistance of the coil in the last problem is 2 ohms, find the total flow of electricity in coulombs, also the average current in ampères, also the energy spent in producing the current.

8. A magnet which includes 6000 lines of force is pulled out of a coil of 160 turns of wire which closely surrounds it, in $\frac{1}{10}$ second. Find the induced electromotive force in volts.

9. A disc of iron 60 cms. in diameter mounted in a uniform magnetic field so that 4000 lines of force per sq. cm. pass perpendicularly through it, rotates like Faraday's disc (§ 722), making 30 revolutions per sec. Find the difference in potential between the edge of the disc and its center.

10. A rectangular loop of wire 20 × 30 cms. is rotated about an axis parallel to the long sides and half-way between them, in a magnetic field of strength 2000. If the axis is perpendicular to the lines of force of the field, what will be the average electromotive force in a half rotation between reversals (§ 743) when the loop makes 20 revolutions per sec.?

11. What is the maximum electromotive force in the case specified in the preceding problem?

ELECTRIC GENERATORS AND MOTORS

Part I. Direct-current Dynamos

742. Introductory. The first machine by which a continuous current of electricity was developed by electromagnetic induction was Faraday's rotating copper disc (§ 722).

A machine developing current by electromagnetic induction consists of a strong magnet between the poles of which an *armature* rotates which contains the conductors in which the currents are induced. Such *generators*, as they are called, are known as *magneto* machines when permanent steel magnets are used, and *dynamo* machines when electromagnets are employed.

743. Rectangular Armature. Suppose that a simple rectangular frame of wire is rotated between the poles of a powerful magnet as shown in figure 428, and that its ends are connected to two rings a and b which are mounted on the axle, and against which press two springs connected to the ends of the outer circuit. In the position shown the upper bar C is rapidly cutting across lines of force. By the rule of induction (§ 732) the induced electromotive force is in the direction of the arrow. So also electromotive force is developed in D. These two electromotive forces act together to cause a current in the outside circuit from B to A. This will be the direction of the current so long as C is moving down across the field of force and D is moving upward. When the coil is in the vertical position both C and D will be moving parallel to the lines of force

FIG. 428. Induction in simple loop armature

so that there is no electromotive force in the circuit at that instant. Then as C comes up and D descends the electromotive force is reversed, causing a current from A toward B in the outer circuit, which reaches a maximum when the coil is horizontal, for then both C and D are cutting perpendicularly across the lines of force. The electromotive force again becomes zero when C reaches the top and D is at the bottom and then reverses again into the original direction.

In the vertical position of the coil the electromotive force is zero, although it includes the maximum number of lines of force, *because in that position a small motion of the coil does not appreciably change the number of lines of force which it embraces.*

While *in the horizontal position the electromotive force is a max-imum*, although no lines of force pass through the coil, because *the change is most rapid in that position.*

The diagram (Fig. 429) exhibits what may be called the curve of electromotive force in such a case. The curve starts
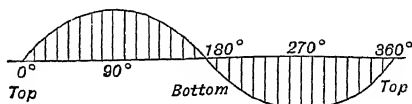
FIG. 429. Diagram of alternating electromotive force

with C at the top, the abscissa at any point is the angle through which C has moved and the corresponding electromotive force is the ordinate, drawn above the horizontal when it is directed from B to A, and below when it is reversed.

The current produced is what is known as an *alternating* current and goes through a complete cycle in the time of one revolution of the armature. *An alternating current may be compared to the surging back and forth of water in a pipe in which a tightly fitting piston is moved to and fro.*

744. Commutator. The terminals of the coil just discussed, instead of being joined to two rings, may be connected to the two halves of a divided ring or *commutator,* as shown in figure 430, on which rest springs or *brushes* which connect to the external circuit and are so placed that they slip from one segment to the other at the instant when the electromotive force in the coil is reversing. In the above diagram, whichever side of the coil is descending, is connected with A, while the ascending side is connected with B, so that the current is always from B to A in the external circuit. The current curve in the external circuit will in such a case be as in figure
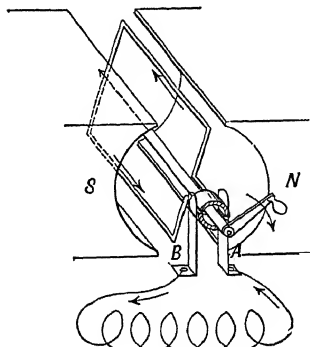
FIG. 430. Loop armature with commutator

431, where ordinates represent the current and abscissas the corresponding instants of time. Each section of the curve

represents half the period of a complete revolution of the armature. Such a current, though always in the same direction, is fluctuating.

745. The Ring Armature. A valuable armature, devised by Pacinotti, is known as the Gramme ring from

FIG. 431

the French inventor who was the first to construct commercial machines using that type of armature. It consists of a soft-iron ring made of a coil of iron wire or a pile of ring-shaped plates of thin sheet iron, wrapped around with a coil of insulated copper wire, the ends of which are joined together forming an endless-ring solenoid with an iron core. For distinctness in the diagram (Fig. 432), the turns of copper wire are shown widely separated. Suppose the ring to be mounted on an axle and rotated between the poles of a powerful magnet as shown in the figure. The lines of force of the magnetic field pass from one pole to the other chiefly through the iron ring as shown by the dotted lines. This, of course, is in consequence of its great *permeability*. As the armature rotates, those parts of the copper winding which cross the outside of the ring cut across lines of force in the space between poles and armature. On the right-hand side the wires cut *down* across
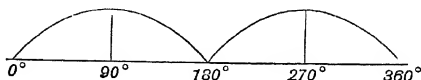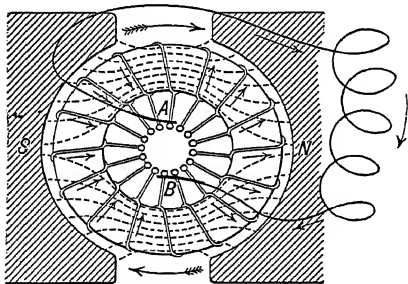
FIG. 432. Gramme ring armature

the field, and the electromotive forces in these turns will be from the front toward the back of the armature. This tends to cause a current in the windings in the direction shown by the small arrows. All of the turns on one side act together like so many little battery cells in series, though those in the middle are most effective. The outer sides of the turns on the left-hand side of the ring, next the south pole of the field magnet, cut up across the field of force, and hence the electromotive force in them is from the back toward the front of the armature, and so they conspire to produce a current on that side in the direction of

the small arrows. But it will be observed that in consequence of the winding of the wire, the induced electromotive force on each side acts to cause a flow around the coil working from the bottom toward the top of the ring, and hence the top of the coil will be a point of high potential and the bottom a point of low potential, when the poles and winding of the armature are as shown in the diagram, but there will be *no flow* around the coil for the electromotive force on one side balances that on the other.

To obtain a current, the top and bottom of the coil must be connected with an outside circuit. This is accomplished by the commutator which consists of a number of segments of copper insulated from each other and mounted in cylindrical form around the axis, each segment being connected with a corresponding point in the copper coil.

The sections of the armature coil included between the points where connection is made to the commutator, all have the same number of turns. In the diagram only one turn is shown for each section, but any number may be used.

If the ends of the external circuit are connected to the two brushes A to B there will be a current from A to B as indicated. For the brush A rests upon the upper segment in the commutator which is connected with the top of the wire coil, and is in this case a point of high potential, while similarly the brush B is in connection with the bottom of the coil where the potential is low.

The flow of current within the armature coil is around on

each side as shown by the arrows, the two currents coming together at the top and flowing out through the commutator at A, around through the external circuit, and in at the bottom of the armature coil where the current divides, half flowing around on one side and half on the other. The case resembles an external circuit connected to

FIG. 433. Two batteries in parallel

two batteries joined in parallel (Fig. 433), the electromotive force of each battery corresponding to that of one side of the armature.

746. Drum Armature. Of every turn of wire on a ring armature part lies on the inside of the ring, and this does not contrib-

ute to the electromotive force. Whatever slight effect it may have, due to the weak magnetic field inside of the ring, is in opposition to the outside part. It is desirable to have as little inactive wire as possible in an armature since it adds to its resistance.

The drum armature is like a ring armature where the opening in the ring is filled up with iron and the turns of copper wire pass clear across the ring from one side to the other, so that the only inactive wire is that across the ends (Fig. 434).

The core is a cylinder of iron made of a pile of thin sheet-iron plates bolted together, around which the coils of wire pass longitudinally lying in grooves made for them. In winding, the wire starts at one of the commutator segments, is passed around the core lengthwise in one of the grooves the desired number of times, suppose twice, and then is connected to the next commutator segment. It is then carried right on around the core in the next groove in the same direction as
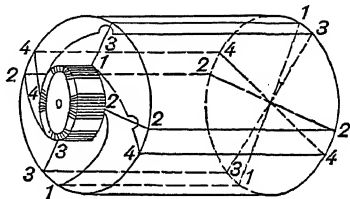
FIG. 434. Drum armature circuit

before, making two more turns, and then connected to the third segment of the commutator. This process is continued until the segment is reached where the winding began and there the end is made fast. In this way an endless coil is constructed just as in the Gramme ring, and between each commutator segment and the one opposite there are two paths by which the current may flow within the armature, so that the current divides in the armature just as in the ring armature.

747. Foucault or Eddy Currents. In each of these armatures the inductive action which causes electromotive force in the copper coils also causes a similar electromotive force in the iron core tending to set up currents within the core itself. Such currents would spend energy in heat, and the double disadvantage would result that more work would have to be spent in turning the armature, and this useless expenditure of energy would go to unduly heat the machine.

In order to prevent these *Foucault currents*, or *eddy currents* as they are often called, the iron core is laminated or made up

of thin plates insulated from each other by varnish, or paper, and lying across the direction in which the currents would flow. The thinner the sheets of iron the more perfectly is this waste of energy prevented.

748. Electromotive Force of Armature. The electromotive force of a ring armature is easily reckoned. *The electromotive force of the ring is the same as that of one side*, since the two sides of the ring act *in parallel*. Let N be the number of lines of force passing through the armature, n the number of revolutions per second, and C the number of turns of wire on the ring, then since each turn cuts down on one side across all N lines of force once in every half revolution, that is in $\dfrac{1}{2n}$ second, the average electromotive force induced in each coil as it moves across the field must be

$$N \div \frac{1}{2n} = 2Nn.$$

But all the coils on one side of the ring act together or in series, hence if there are C coils of wire on the ring the total electromotive force must be

$$2Nn \cdot \frac{C}{2},$$

thus

$$E = NnC \quad \text{in C. G. S. units,} \quad \text{or} \quad E = \frac{NnC}{10^8} \text{ volts.}$$

The electromotive force depends on three factors: the number of lines of force through the armature, the number of revolutions which it makes per second, and the number of coils of wire upon it.

The electromotive force of a drum armature is calculated from the same formula, C representing the whole number of wires on the armature which cut across lines of force.

749. Field Magnets. In most generators and motors the armature rotates between the poles of an electromagnet which receives its exciting current from the armature. Three modes of winding are in use, *series, shunt,* and *compound.* In the

first diagram in the figure is shown a *series*-wound generator. The whole armature current passes around the field magnets and through the external circuit. Any resistance introduced into the external circuit, causing the current to diminish, weakens the magnetic field and therefore makes the electromotive force of the machine less. When there is no current flowing its electromotive force is zero except for the residual magnetism.

In the *shunt* arrangement the current in the armature divides, part flowing around the magnet and part to the external circuit. In order that but a small current may be taken for the magnet. it is wound with many turns of rather fine wire.
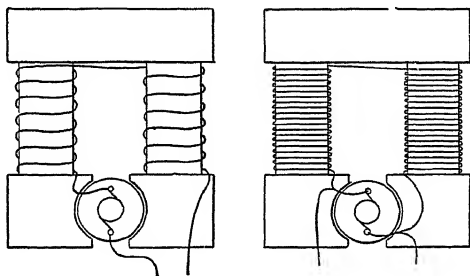
FIG. 435. Series and shunt field magnets

The current through the shunt coil depends only on its resistance and on the difference of potential of the brushes; hence it is constant and the strength of the magnet is constant so long as the difference in potential of the brushes is unchanged. *The electromotive force of such a dynamo is very nearly constant*, but is slightly greater when no external current is flowing, for with increasing current in the external circuit there is more current and a greater *fall of potential in the armature itself*.

Compound winding is a combination of the shunt and series arrangements, in which there is a shunt coil and also a few turns carrying the whole current around the magnets. In this way a generator may be made to maintain a nearly constant potential at the terminals, though the external current may vary greatly, or it may be *over*-compounded so that its terminal electromotive force may be greater with large currents than with small.

Part II. Direct-current Motors

750. Motors. An electric motor is an appliance in which an electric current gives motion to an armature, thus producing mechanical work. Small direct-current motors usually have ring or drum armatures and are in most respects like generators.

The action of the ring armature in a motor may be understood from the diagram (Fig. 436). The current from a battery or other source is shown as flowing in at the upper brush and out at the lower one. Within the armature the current divides, half flowing around and down through the coils on one side and half through those on the other side as shown by the arrows, and the effect of these currents in the armature is to make each half of the ring a magnet with its north pole at the top and south pole at the bottom. The attractions and repulsions between these poles and those of the field magnet cause the armature to rotate in the direction of the large arrows.

FIG. 436. Motor with ring armature

Another aspect of the action is worth considering. The gaps between the pole pieces and the armature are regions of intense magnetic force, and the wires on the outside of the armature carry currents directly across these lines of force, *up* (perpendicular to the paper) on the left and *down* on the right; there is, therefore, a force (§ 697) urging these wires to move across the lines of force toward the top of the diagram on the left and toward the bottom on the right.

751. Energy Spent in Motor. While the motor is running mechanical work is being done in addition to the energy which is spent as heat in the armature in consequence of its resistance. But the total energy spent per second in the motor is equal to the product of the current strength by the difference of potential between the brushes. Therefore if the current is kept constant the difference of potential between the brushes must be greater when the motor is running and doing work than when the armature is at rest.

This increase in the difference of potential between the brushes due to the

motion of the armature is *the back electromotive force* of the motor. *There must be such a back electromotive force in every kind of device in which motion results from the flow of an electric current.*

Let $V_1 - V_2$ = difference of potential between brushes of motor.

IR = drop in potential due to the resistance of the armature.

$V_1 - V_2 = E + IR$ where E is the back electromotive force.

Total watts spent in motor = $\begin{cases} \text{watts spent in turning armature} + \text{watts} \\ \text{spent in heat} \end{cases}$

or in symbols

$$I(V_1 - V_2) = IE + I^2R.$$

752. Back Electromotive Force. Connect an electric motor to a battery by which it may be driven and introduce into the circuit an incandescent lamp which will glow with full brilliancy when the armature of the motor is held stationary. *On letting*
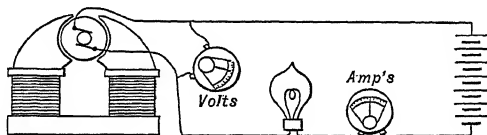
FIG. 437

the armature run the lamp grows dim, and an ammeter in circuit shows that *the current has diminished,* but *a voltmeter connected to the brushes of the motor will show a much greater difference of potential between them than when the armature was at rest.*

Since the electromotive force of the battery and the resistance of the whole circuit is unchanged by the running of the motor, it is clear that *the current can have been diminished only by the development of an electromotive force in the circuit back against the driving current.* The motor, in fact, while running acts like a dynamo and develops an electromotive force, called its *back electromotive force,* because it acts in opposition to the electromotive force of the driving battery.

753. Starting a Motor. In starting a motor there is at first no back electromotive force to oppose the current, and in order to prevent the current being excessive and "burning out" the armature before the motor is well started some such device as shown in figure 438 is commonly used.

The current is led to the motor through the wires AB, one

of which is connected directly to the motor while the other is joined to the switch S. When the switch is turned from 1 to 2 the current flows through coils of wire having considerable resistance and starts the armature.
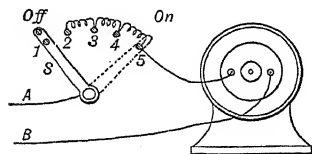
FIG. 438. Starting connections of motor

As the speed increases, developing more back electromotive force, the switch is moved on to 3 and 4, reducing with each step the extra resistance until, as the armature comes to full speed, the switch on 5 makes direct connection, and the back electromotive force keeps the current moderate even though the armature resistance may be extremely small.

754. Generator and Motor. Suppose a transmission system consisting of generator and motor and connecting circuit. Let the electromotive force of the generator be 200 volts, and suppose the resistance of the whole circuit including the armatures of both generator and motor to be 1 ohm, and let the back electromotive force of the motor be 180 volts at the working speed.
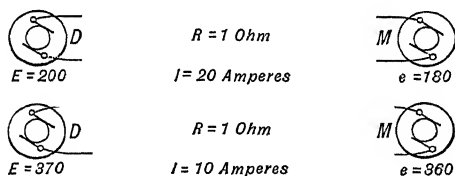
FIG. 439

Then the resultant or effective electromotive force in the circuit is $200 - 180 = 20$ volts, and the current is 20 ampères.

Power spent in the generator	$200 \times 20 =$	4000 watts.
Power used in motor	$180 \times 20 =$	3600 watts.
Loss in heat (I^2R) is the difference		400 watts.

If a motor is used which in running develops twice the back electromotive force of that just discussed, then with a current of 10 ampères as much power will be obtained as with the 20 ampères in the former case.

In this case the electromotive force of the generator must be

370 volts, and that of the motor being 360 volts the effective electromotive force is 370 − 360 = 10 volts. The current will therefore be 10 ampères, and we have

Power spent by generator = 370 × 10 = 3700 watts.
Power used in motor = 360 × 10 = 3600 watts.
Power wasted in heat I^2R = 100 × 1 = 100 watts.

This is evidently a much more economical arrangement than the first and illustrates the general principle that electrical energy can be transmitted with least loss by means of small currents at high voltage.

Part III. Alternating Currents

755. Alternating Currents. Alternating currents have come extensively into use because of the ease with which a large alternating current at a low voltage can be changed to a small one at a high voltage. The small high-voltage current can be carried by comparatively small conductors to a distant point and then be transformed down again to a large current at a low enough voltage to be safely used for light or power.

756. Alternating-current Generator. Almost any direct-current generator will give alternating currents if it is provided with two rings mounted on the axis and connected respectively to two diametrically opposite segments of the commutator. A circuit whose ends are connected to these rings by brushes will have an alternating current. Such a case was illustrated in § 743. To secure good insulation, high electromotive force, and sufficient frequency of alternation, alternating-current generators are usually multipolar, as illustrated in figure 440. In the type shown the field magnet poles, alternately north and south, project outward from the rim of a rotating wheel and are magnetized by the current supplied by a small separate direct-current genera-tor called an exciter. This rotary field, or rotor, rotates within the fixed armature or stator, in which the poles project inward from the outside circular frame. These poles are of the same number as those on the rotor, and are laminated or built up of thin plates of sheet iron to prevent eddy currents. Around the poles of the stator the armature coils are fitted, passing through

slots between them and wound alternately clockwise and counter-
clockwise around successive poles. As the rotor turns and a north
pole facing one pole of the armature moves over to the next, the
lines of force from the pole of the rotor cut across the conducting
wires lying between the poles of the stator and induce electro-
motive force in them which reverses as the succeeding south
pole moves across, thus causing an alternation. But as the
wires in one slot return in the opposite direction through the
next slot, and as a north pole is moving across the one while a
south pole is moving across the other, the electromotive forces
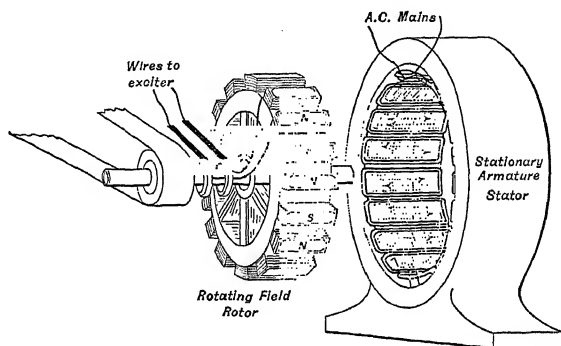
FIG. 440

induced in all act together at every instant, so that in the case
figured where there are 16 poles, an alternating electromotive
force is produced 16 times as great as would be developed in a
single coil.

If a low electromotive force is desired the several coils of the
armature may be connected in parallel instead of in series as
above described.

757. Virtual or R. M. S. Ampères and Volts. An alternating
current is constantly varying in strength, as illustrated in the
curve of figure 441, its average value is zero and it will not give
a steady deflection of the needle in an ordinary galvanometer.
A definition must therefore be given of what is meant by an al-
ternating current of one ampère. Since the energy relations of a
current are commercially the most important, *an alternating cur-*
rent is said to have the strength of 1 *ampère, when it will develop the*

same amount of heat in a given resistance as would be produced by a direct current of 1 ampère. The heating effect of a current at any instant is proportional to the square of its strength at that instant, so also the deflection produced by a current in an electrodynamometer is proportional to the square of the current strength (§ 705); therefore an electrodynamometer measures directly the virtual ampères of an alternating current just as it does a direct current.

If the alternating-current curve is a *sine* curve, the virtual strength as defined above is to its maximum value in the ratio of 1 to 1.41 (1.41 = $\sqrt{2}$); thus an alternating current of 10 ampères ranges from +14.1 to −14.1 ampères in its instantaneous values.
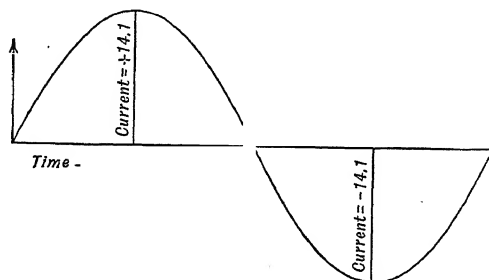
FIG. 441. Alternating-current curve. Current = 10 Ampères

So also the virtual value of an alternating electromotive force is said to be 1 volt when it will develop an alternating current of 1 ampère in a resistance of 1 ohm having no self-induction or inductance. It is common practice to use the term " R. M. S." (meaning "Root Mean Square") in place of the term "*virtual.*"

758. Effect of Inductance. It has already been shown (§ 733) that the effect of inductance in a circuit is to cause an electromotive force *contrary* to an increasing current and *with* a decreasing current. In case of alternating currents, the effect is twofold. *First, it causes an apparent increase in resistance.* It may be proved that the current produced by an alternating electromotive force E in a coil whose coefficient of inductance is L and whose resistance is R, is

$$I = \frac{E}{\sqrt{R^2 + (2\pi f L)^2}}$$

where f is the number of complete cycles per second, or the *frequency*, the current and electromotive force being measured in virtual ampères and volts, the resistance in ohms and the inductance in henrys. The denominator is known as the *impedance* of the conductor.

If the inductance of the coil is large and if there are a large number of alternations per second, the impedance may be large although the resistance is small.

Second, inductance causes the phase of the current to lag behind that of the electromotive force, so that the current does not reach its maximum value at the same instant that the electromotive force is a maximum, but a certain fractional part of a period later, which is called the lag.
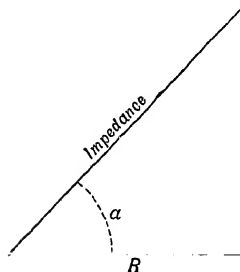
FIG. 442

The relations of these quantities are shown by the triangle in figure 442. If the base of the right-angled triangle represents the resistance of a coil, and the altitude, the quantity $2\pi fL$, then the hypothenuse represents the impedance, and the angle α at the base, the angle of lag. That is, the current maximum lags behind the maximum of electromotive force the same fractional part of a complete period that α is of the whole angle about a point.

759. Theater Dimmers. If an electric glow lamp, connected in series with a coil of wire having very low resistance, is lighted by means of an alternating current, the light may be dimmed by inserting a laminated core of soft iron inside the coil. The inductance of the coil is greatly increased in this way and the current is decreased, but there is no waste of energy as there would have been if the current had been reduced by introducing resistance. This method is used for dimming theater lights.

760. Transformers. *Alternating currents are easily changed from low voltage to high, or vice versa, by means of transformers.* A transformer consists of two coils side by side, having a common core of soft iron. In the form shown in figure 443, the iron core is made up of a pile of thin sheet-iron plates of the shape shown in the section. The core thus formed is a block of soft iron having

two rectangular holes through it in which the two coils lie side by side, one coil having many turns of fine wire and the other a few turns of coarse wire as shown in the diagram. When an alternating current is set up in one coil it magnetizes the iron core, setting up lines of force which at one instant are in the direction shown by the arrows in the diagram and a half period later are exactly opposite. But the lines of force pass through the second coil as well as the first and therefore an alternating induced current is set up in the secondary coil.
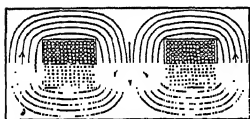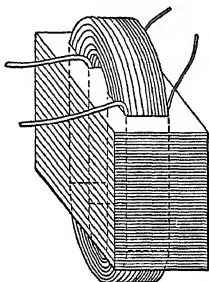
FIG. 443. Transformer and section   FIG. 444. Transformer connections
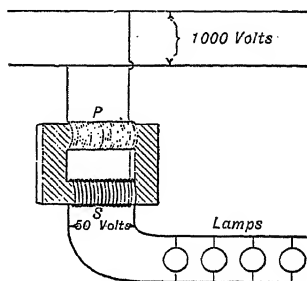
The same lines of force cut across one coil as the other and consequently the electromotive force induced in one coil is to that in the other as the number of turns of wire in the coils.

Suppose it is required to transform from 1000 volts down to 50. The fine wire coil which is connected to the 1000-volt circuit must have 20 times as many turns of wire as the coarse wire coil which is connected with the lamps. If no lamps are turned on, there is no current in the secondary coil and the magnetic field through the primary coil causes such a strong back electromotive force that only a very small current flows through it and there is but a small loss of energy.

When lamps are turned on in the secondary a current flows which by the laws of induction tends to oppose the changes in

flux produced by the primary magnetizing current. This latter current then rises to a value sufficient to sustain the original alternating flux necessary to maintain the back electromotive force which balances the electromotive force of the main line.

In this way the transformer is self-regulating, the primary current being very nearly in the same ratio to the secondary as the number of turns of wire in the secondary coil is to that in the primary. In the example considered, the current in the primary would be one-twentieth that in the secondary.

The energy spent in the secondary circuit is equal to that which the transformer takes from the main line except for a small amount, 2 or 3 per cent, which is lost as heat in the transformer.

761. Advantage of Transformers. Large currents cannot be transmitted long distances without great loss in heat unless large conductors of low resistance are used, in which case the cost and interest charges are high. By means of a transformer a large electrical power may be transmitted by a small current at high voltage. Thus in districts where scattered houses are to be lighted a small current at high voltage is used on the street line and transformed down, giving large currents at low potentials at the points where lights are used.

In many lines where power is to be transmitted a long distance transformers are used at both ends of the line. Thus at Niagara generators develop currents at 13,800 volts, which are then transformed *up* to 66,000 volts, and so transmitted to Buffalo, and to other cities from 20 to over 100 miles away, where they are transformed down again for power and lighting purposes.

762. Electric Welding. An important application of large electric currents is in fusing bars of metal together. Two bars of iron as large even as a man's wrist may be placed end to end and fused together in a few seconds. For such a purpose a very large current is required just at the spot to be heated. Accordingly a transformer is used in which the secondary may consist of only a single turn or two built of heavy copper bars, the terminals of which are clamped to the bars to be welded, one on each side of the junction. The primary coil is made of many turns of wire and takes a comparatively small current at high voltage.

763. Alternating-current Motor. There are two principal

types of alternating-current motors, the *synchronous motor* in which the armature will not start of itself, but must be brought by some accessory motor to such a speed that its armature coils move from one field pole to the next in exact synchronism with the alternations of the driving current. When brought to speed, it will continue to work when driven by a single alternating current.

A second type is the *induction motor* in which a rotary magnetic field is produced by polyphase currents (§ 765).

764. Rotary Magnetic Field. Suppose that a laminated ring of soft iron, having four poles projecting inward as shown in figure 445, is wound with two in-dependent circuits, one of which magnetizes the *A* and *C* poles and the other the *B* and *D* poles. And let an alternating current be es-tablished in each circuit, the phase of the current in the *E* circuit being a quarter of a period ahead of that in the *F* circuit, as shown in the curves *E* and *F* in figure 446, so that one reaches its maximum value, either posi-tive or negative, at the instant that the other is passing through its zero value.
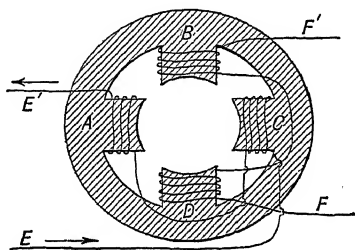
FIG. 445. Rotary field magnet

The corresponding changes in the direction of the lines of force in the field between the poles are shown in the lower dia-grams of figure 446. Thus in the first diagram the current *E* is a maximum, and is supposed flowing from *E* to *E'* (Fig. 445) mak-ing *A* a north pole and *C* a south pole, while at that instant the *F* current is zero. But as the *F* current increases that in *E* dimin-ishes until *F* becomes a maximum and *E* zero. The north pole has now passed to *B*, while *A* and *C* have lost their polarity as shown in diagram 3. The sign of the *E* current is now re-versed and it begins to flow from *E'* toward *E*, making *C* a north pole as shown in 4; at the same time the *F* current is decreasing and becomes zero in 5, where *E* reaches its maximum negative value. In this way what is known as a *rotary magnetic field* is produced in which the north and south poles move around the ring making one revolution for every complete period of the current.

765. Induction Motors. Between the poles of the rotary field just described there is mounted a cylindrical-shaped armature having a set of parallel rods of copper at equal intervals around the circumference, like the bars in the wheel of a squirrel cage, connected across the ends by copper plates. And to strengthen the lines of force through the armature, it is filled
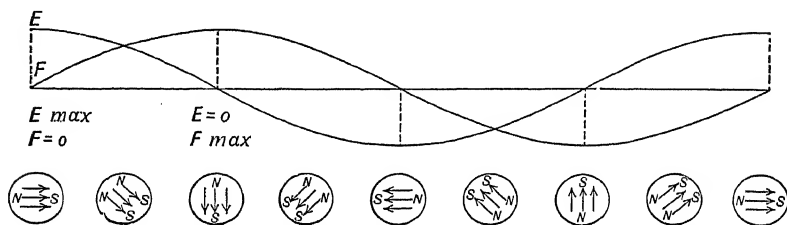
FIG. 446.   Diagram of two phase currents and rotary field

with a soft-iron core made of a pile of circular plates of thin sheet iron. As the lines of force of the field rotate they cut across the bars of the armature, inducing currents which by Lenz's law are in such a direction as to resist the relative motion of armature and field, and the armature is therefore carried around in the direction in which the field rotates. But clearly in such an *induction motor, the armature cannot rotate as fast as the magnetic field,* because it is the difference between the motions of the two that causes the induction on which the rotation of the armature depends.
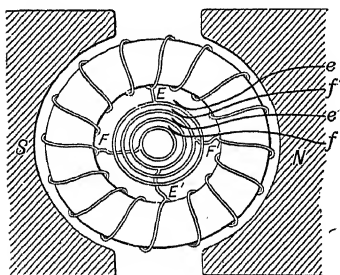
FIG. 447.   Connections for currents in quadrature

766. How Currents in Different Phases are Obtained. Imagine a Gramme ring armature as shown in figure 447 provided with four insulated brass rings mounted on its axis, each of which is connected permanently to one of the points $EFE'F'$, which are just one-quarter circumference apart on the ring. If one circuit is now connected to the brushes e and e' and another to the brushes f and f' which rest on the rings, the currents in the two circuits will be one-quarter period different in phase as represented in figure 446.

767. Three-phase Motors. The usual form of induction motor uses *three-phase currents,* or three currents which differ in phase by one-third of a period, and requires only three line wires instead of four. The generator has three rings connected, re-
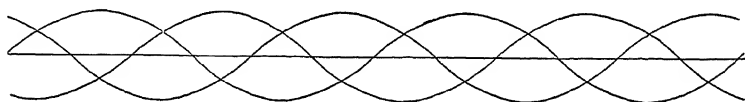
Fig. 448.   Three-phase currents

spectively, to three equidistant points in the armature, so that the currents developed in the three line wires are related as shown in the curves of figure 448. It will be noted that the sum of the ordinates of any two of the three curves taken at any point along the base is equal and opposite to the ordinate of the third curve at that point; that is, *the sum of the currents in any two of the*
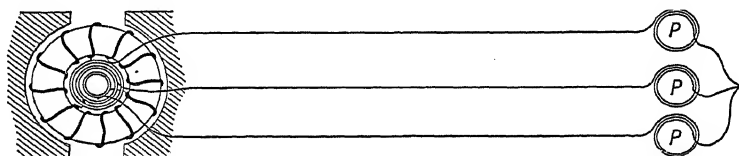
Fig. 449.   Connections of three-phase generator to the poles P
of the field magnet of the motor

three line wires at any instant is equal and opposite to the current in the remaining line wire; the three are, therefore, connected together at the farther end and *each serves as the return wire for the other two,* as shown in figure 449.

Three-phase motors are usually multi-polar, each principal pole being subdivided into three parts. The figure shows a field having twelve small poles which are so wound as to form a rotary field with two north poles and two south poles. How this is done may be understood from the diagram in which the field ring is supposed to be cut

Fig. 450

at one point and bent out flat so that we look directly at the faces of the twelve poles.
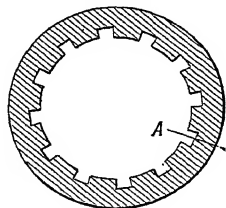
For simplicity the wire is represented as carried only once

through each groove. It will be seen that when the current in 1 is a maximum in the direction of the arrow the poles will be situated as shown in the upper row of letters. A third of a period later the current in 2 will be a maximum in the same direction, and the poles will then be as indicated in the second row. Then after another one-third of a period current 3 will have reached its maximum and the poles will have shifted to the positions indicated in the third row of letters. There is thus produced a
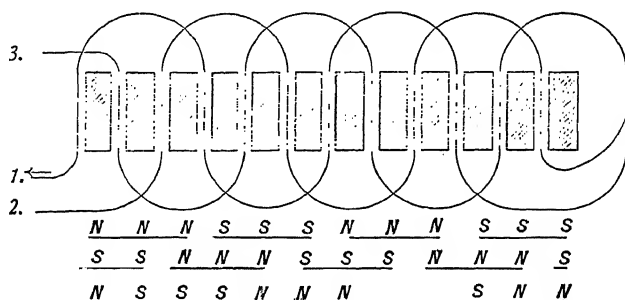
FIG. 451. Windings of a three-phase field magnet

steady movement of the poles around the ring, moving over the distance between two similar poles in the time of one complete period of alternation of the current. In the above case, if the current has a frequency of sixty periods per second, the field will make thirty revolutions per second.

PROBLEMS

1. The core of a Gramme ring armature has a cross section of 6 × 10 cms. How many turns of wire must it have that it may give an electromotive force of 20 volts when making 800 revolutions per minute in a magnetic field so strong that where the lines of force in the ring are most concentrated there are 6000 per square centimeter?

2. A certain dynamo armature when making 1000 revolutions per minute is supplying a current of 50 ampères at 100 volts. Find the horse-power required to drive it and thence the moment of force or torque in pound-feet required to turn the armature at the given speed.

3. When the armature of a certain motor is held fixed a current of 10 ampères through it causes a difference in potential between its brushes of 5 volts. When the armature is permitted to run at 600 revolutions per minute the current is 4 ampères and difference of potentials at the brushes is 30 volts. Determine the back electromotive force of the motor.

4. The core of a drum armature is a cylinder of iron 30 cms. long and 15 cms. in diameter, the induction through its middle longitudinal section is 6000 lines of force per square centimeter. If there are 50 complete turns of wire on the armature, or 100 longitudinal wires in grooves on its surface, what is its electromotive force when making 1200 revolutions per minute?

5. A transformer has a coil of 250 turns; what must be the size of the iron core in order that an average alternating electromotive force of 100 volts may be developed in this coil while the number of lines of force in the core changes from $+6000$ to -6000 per sq. cm., the current alternating at the rate of 60 complete periods or cycles per second?

6. A certain transmission line has a resistance of 20 ohms. How much power will be lost in the line when 100 kilowatts are transmitted at 2000 volts? How much when the same power is transmitted at 20,000 volts?

7. A multipolar generator having 16 poles (Fig. 440) makes an alternating current of 60 cycles per sec. How fast does it rotate? If there are 30 turns in each armature coil, what E.M.F. is developed when each pole of the rotor gives rise to 100,000 lines of force?

Electric Discharge through Gases

768. Nature of Gaseous Conduction. At ordinary temperatures and atmospheric pressure gases are almost complete non-conductors of electricity. Careful experiments have shown, however, that there is a slight loss of charge through the air, which is due to the presence in the air of a few *ionized* molecules, or molecules which are broken up each into a positive ion and a negative ion or electron.

The electric current is transmitted through the gas at low pressure by means of the positive and negative ions, which act as carriers. With no ions, no current can pass. The same is true of conduction through liquids (§ 627). In a gas, however, the negative ion is often the electron itself.

769. Causes of Gaseous Ionization. Since a gas cannot conduct electricity unless ionized, the question arises, how do gases, which conduct electricity, become ionized? Two important causes of ionization are collisions between electrons and molecules and the direct action of radiation.

In the former, the few free electrons always present in a gas are acted upon by the applied electric field and if the field is strong enough these electrons are driven against neutral molecules with sufficient force to knock off loosely held electrons, thus producing positive ions and electrons. These new electrons in turn produce

more by the same action, the effect being cumulative so that the number of ions may multiply very greatly in a short distance. This cumulative production of ions by collisions is an important cause of electric discharge through gases.

Ionization by radiation is observed readily by exposing a mass of gas to X-rays (§ 782). The gas becomes conducting immediately so that a gold leaf electroscope quickly discharges. X-rays and ultraviolet light (§ 934) have greater power to ionize than radiation of longer wave lengths. The ionization of gases may be brought about by chemical action and molecular activity at high temperatures so that flame gases readily conduct electricity.

770. Electron Emission from Solids. Another important consideration in gaseous conduction is, how the electrons, which make up the conduction current, are removed from the negative electrode. A continuous supply of them is necessary, which can hardly be accounted for without their being directly pulled out of or emitted from this electrode. The electrons within a conductor may be thought of as held there as the molecules are held within a mass of liquid. They move easily from one part of the liquid to another, but are removed from the surface only by expending a certain amount of energy. At high temperatures, when the molecules are greatly agitated, the attractive force is overcome, and they fly off in the form of a vapor (§ 442). Exactly the same thing is true of the electrons within a conductor. If the temperature is sufficiently high, they are thrown off as a vapor of electrons. In fact evaporation from a liquid and the emission of electrons produced by high temperature are governed by very similar laws.

771. Discharge at Atmospheric Pressure. The difference in potential between two knobs required to cause a spark to pass between them at atmospheric pressure appears to be nearly the same whatever metal is used for the knobs, but it depends on their curvature.

For knobs over 2 cms. in diameter and more than 2 mm. apart, the number of volts required for spark discharge is approximately given by the formula

$$V = 30,000d + 1500$$

where d is the distance between the knobs in centimeters.

TABLE OF SPARK POTENTIALS BETWEEN SLIGHTLY
CURVED SURFACES IN AIR

Spark Length	Voltage	Voltage per Centimeter
0.0015 cm.	426	284,000
0.01	948	94,800
0.1	4,419	44,190
0.5	16,326	32,652
0.8	25,458	31,822
1.0	31,650	31,650

Heydweiler

772. Effect of Diminished Pressure. On diminishing the
pressure, the potential necessary for discharge becomes lower
until a certain critical degree of exhaustion is reached. Beyond
this point the higher the exhaustion, the greater the potential dif-
ference required to produce discharge, until at the highest ex-
haustions a spark can hardly be made to pass, and discharge will
take place through several inches of air at atmospheric pressure,
in preference to 1 mm. in the vacuum.   The critical pressure
is less than a millimeter of mercury when the electrodes are more
than a few millimeters apart, but if they are very close it is
somewhat greater.

The appearance of the discharge may be conveniently studied
in a wide glass tube 3 or 4 ft. in length, closed at the ends with
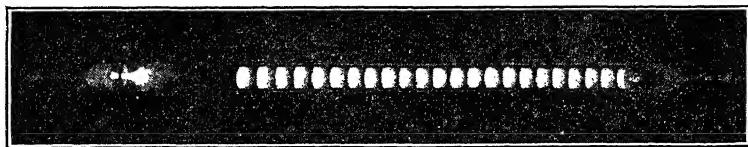
Fig. 452. Discharge in gas at low pressure

caps in which the electrodes are mounted, and connected
with an air-pump. If the electrodes are connected with an in-
duction coil or electrical machine, the discharge will take place
through the tube after a few strokes of the air-pump. At first
there is a crackling, flashing discharge along narrow flickering
lines, but as the exhaustion proceeds the lines of discharge widen
out and fill the whole tube, which glows with a steady light.

The discharge at first is between certain points on the electrodes, but with higher exhaustion the luminous glow entirely covers the surface of the negative electrode.

In this stage the characteristic features of the discharge are as follows: A faint velvety glow covers the surface of the negative electrode or cathode; just outside of this is the Crookes dark space which surrounds the cathode and has nearly a constant width everywhere. Then comes a luminous region, called the negative glow, and then a dark space, the so-called Faraday dark space, after which a luminous column, known as the positive column, reaches all the way to the anode.

The positive column is commonly not continuous, but shows alternate bright and dark layers across the path of discharge; and these bright layers or *striæ* become broader and farther apart as the exhaustion is increased.

If the distance between electrodes is increased the appearance at the negative electrode is not particularly changed, the positive column, however, is increased in length and reaches as before nearly to the negative glow. Professor J. J. Thomson examined the discharge in an exhausted tube 50 ft. long, and found that the positive column reached the entire length of the tube to within a short distance of the negative electrode and was stratified throughout.

773. Explanation of Effects Observed at Diminished Pressure. Electric discharges through gases are doubtless first started by the building up of ionization by collision, which so multiplies the number of ions and free electrons that the gas within the tube becomes conducting, due to the free flow of electrons from the cathode to the anode. At the same time positive ions near the negative electrode are attracted to it.

When the pressure of a gas is lowered, the distance between the molecules becomes greater so that the few ions always present will move greater distances under the electrical force and thus acquire higher velocities before striking molecules. Thus a smaller difference in potential than before will be sufficient to cause the building up of ionization and start an electrical discharge. At very low pressures, when the molecules travel several cms. between collisions electrical discharge takes place at very low voltages, which may be not more than 150 volts. In such

cases the discharge is probably maintained by electrons liberated from the cathode by impacts from the positive ions which continually bombard it.

At higher evacuation ionization by collision begins to fail because the electrons intercept fewer and fewer molecules before hitting the anode owing to the greater scarcity of the molecules whose presence make the collisions possible. Therefore the means of starting and maintaining the discharge begins to fail as the voltage rises until, at the highest exhaustion, no discharge at all can be made to pass even at the highest voltages.

774. Geissler Tubes. Geissler tubes, so called from the name of a well-known maker who showed great skill and ingenuity in their construction, are tubes of glass especially designed for the purpose of exhibiting the phenomena of discharge. They are exhausted to a pressure of about 1 mm. of mercury, and are provided with aluminum electrodes attached to wires sealed into the glass. They are usually wide near the electrodes, but often a part of the tube is quite narrow, and here the concentration of the discharge makes the illumination particularly brilliant and the stratification very noticeable.

In some tubes a marked fluorescence of the glass is produced by the discharge, some kinds of glass glowing with a yellowish green light, while other kinds appear bluish. When such a tube is surrounded by a solution of sulphate of quinia or fluorescein or other fluorescent liquid, the characteristic fluorescence is strikingly brought out.

The character of the light from such a tube depends on the gas which it contains, a tube containing nitrogen or atmospheric air appears of a reddish-violet color, a hydrogen tube is much bluer, while carbon dioxide gas shows a pale whitish illumination. The light of each when analyzed by a spectroscope is found to be made up of certain particular wave lengths characteristic of the gas.

775. Cathode Rays. In tubes exhausted considerably beyond the pressure of greatest conductivity, say to about one-thousandth of a millimeter of mercury, where a considerable voltage must be applied to cause a discharge of electricity through the tube, the electric discharge consists primarily of a stream of electrons which pass through the vacuum at high velocity from

cathode to anode. This stream of electrons, which is called the *cathode rays*, is projected out nearly at right angles to the surface of the cathode, without reference to the position of the positive electrode, as would be expected of a stream of electrons moving

at high velocity. In the tube shown in Fig. 453, a sharply defined shadow of the metal cross is cast on the end of the tube opposite the cathode, for the rays excite brilliant yellowish fluorescence in the glass wherever they strike it directly. The position of the positive electrode is immaterial. A crystal of Iceland spar or calcite on which the cathode rays fall glows with orange red light which persists

FIG. 453. Crooke's tube with screen intercepting cathode rays

some seconds after the discharge has ceased.

If the cathode is concave, the electrons may be concentrated onto a small arc on a surface and produce intense heat which may be sufficient to melt even platinum.

776. Nature of Cathode Rays. That the cathode rays are actually a stream of high velocity electrons projected from the cathode has been proved beyond doubt experimentally. The following are two of the tests which have been used to establish this.

In the tube shown in figure 454 the rays from the cathode after passing through a narrow opening in the screen S fall upon a sheet of mica running lengthwise with the tube and covered with fluorescent material so that the path of the rays is distinctly seen. The boundaries of the luminous path are seen to curve outward slightly as though the discharge consisted of a stream of particles charged with electricity whose mutual repulsion causes them to separate while they are streaming forward.

If a horseshoe magnet is now held with its poles on opposite sides of the tube so that its lines of force are at right angles to the path of discharge, the rays are deflected to one side. If the lines of force are down, perpendicular to the paper, the rays will be bent into the position shown by the shaded curve, just as would be expected of particles of matter negatively charged and projected forward.

It has also been shown by J. J. Thomson that the stream of cathode rays is deflected when a positively or negatively charged body is brought near it, being attracted by the former and repelled by the latter.

From the amount of deflection of the rays in a magnetic field of known strength, taken in connection with the deflection caused by a known electrostatic field, J. J. Thomson estimated that the particles forming the cathode rays have each a mass of about $\frac{1}{1840}$ that of a hydrogen atom, and move with a velocity which
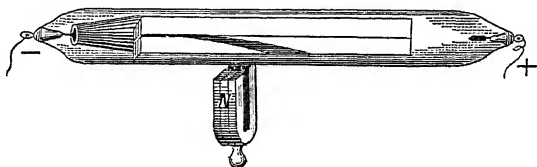
Fig. 454. Cathode rays deflected by magnet

depends only upon the fall of potential in the tube, and may be about $\frac{1}{10}$ of the velocity of light or about 18,000 miles per second. Each particle carries a negative charge equal to 4.77×10^{-10} electrostatic units.

By such experiments Sir J. J. Thomson about the year 1898 was the first to show that electricity is not a fluid but is atomic in its structure, and to measure the mass and charge of this atom of electricity known as the electron (contraction for electric ion).

777. Hot Cathode Tubes. It has been mentioned already that if a tube is sufficiently highly evacuated no electrical discharge can be made to pass through it, however much the voltage between the electrodes of the tube is increased. If the cathode be heated to a high temperature, however, electron emission takes place directly as a result of thermal agitation. To accomplish this, the cathode is made in the form of a filament, usually of tungsten, so that it can be heated by passing an electric current through it. As mentioned in § 770 the thermal emission of electrons is governed by laws similar to the evaporation of molecules from a liquid surface. The rate of emission depends upon the temperature according to a definite law. Figure 455 shows the electron current emitted per square cm. from the surface of tungsten for different temperatures. Above red heat

the rate of emission rises extremely rapidly indeed until limited by *space charge* (§ 780). If the temperature is raised above 3000° C. the tungsten metal itself begins to evaporate an appreciable amount so that the life of the filament is decreased.
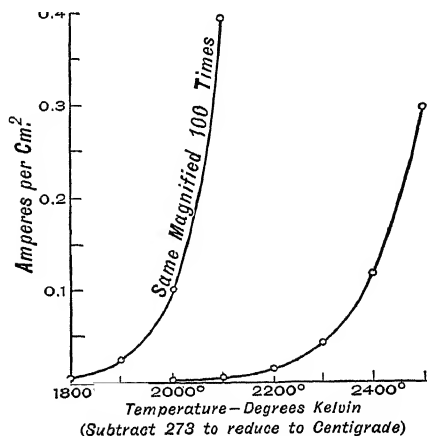
FIG. 455. Thermionic emission from tungsten

778. The Work Function. The affinity of a surface for electrons is different for different substances. It is measured by what is called the *work function* of that surface which is proportional to the work necessary to pull an electron out of the surface. The electrons within the surface are moving with all sorts of different velocities. The fastest moving electrons shoot out from the surface against its attractive force, the number which escape in this way per second rapidly rising with temperature. This accounts for the rising curve of figure 455. In general the smaller the work function of a substance, the lower the temperature required to produce a given electron emission. Therefore many hot cathode tubes are used with their filaments coated with oxide of calcium, barium, strontium or caesium, whose work functions are of the order of half that of tungsten or of platinum. These filaments operate at much lower temperatures than non-coated tungsten filaments.

779. Thermionic Rectifiers. An important application of the emission of electrons from filaments has been made in *thermionic rectifiers*. A rectifier is a device used to *rectify* an alternating electric current so that it flows in one direction only, thus becoming a direct current.

The alternating voltage terminals are connected to the primary of a transformer as shown in figure 456 and the rectifier is

FIG. 456. Thermionic rectifier

connected in the secondary circuit. It acts as a valve in this circuit permitting a current of electrons to flow from the filament across the tube to the opposite electrode, but not in the reverse direction.

The induced alternating electromotive force in the secondary coil causes the potential of the filament alternately to fall below and rise above that of the opposite electrode once every cycle. When the filament voltage is low, or when it is negative, electrons can leave the filament and cross to the opposite electrode. When the latter is negative, however, the current is completely cut off since there are no carriers present to carry charge in the opposite direction. The thermionic rectifier thus acts as a valve in the secondary circuit, allowing a pulse of current to flow *out* of the filament only, once every half cycle. The resultant effect is a pulsating current which always flows in the same direction.

The figure shows a tungar rectifier which is being used to charge a storage battery. This type of thermionic rectifier is designed for low voltages and comparatively large currents. The current is far larger than is possible from thermionic emission alone. The result is brought about by the presence of a small amount of the inert gas argon which becomes ionized by collisions, the positive ions also taking part in producing current. If the alternating voltage across this type of tube be increased beyond a certain point, current flows through the tube in both directions, and its rectifying action breaks down. Low voltages only can be used on a thermionic rectifier containing gas.

Another type of thermionic rectifier called the *kenotron* is designed for use in applications where high voltage is required. This tube is highly exhausted so that the so-called thermionic current only is made use of. It is operated through a transformer just like the tungar rectifier. By its use direct currents of over 100,000 volts may be produced, but they are small, not over $\frac{1}{4}$ to $\frac{1}{2}$ ampères, since they are limited by the magnitude of the thermionic current which can be emitted from the filament.

780. Space Charge. If the voltage across a thermionic vacuum tube with two electrodes, like the kenotron, be raised gradually, the electron current rises gradually also as shown in figure 457 until a certain voltage is reached beyond which the current changes but slightly. The latter effect is easily explained since at a particular filament temperature electrons are emitted at a definite rate, and an increase of voltage has little effect upon the number emitted per second, characteristic of the filament temperature.

It is very surprising indeed, however, that current curve of figure 457 gradually falls off below a certain voltage. This shows that unless the voltage is sufficiently high something prevents the electrons from leaving the filament as fast as they would if they were free to escape. This was proved by C. D. Child and I. Langmuir to be caused by the repulsive action of the swarm of electrons, which are passing through the tube, upon those at the surface of the filament. Each electron as it is emitted, although attracted by the positive electrode, is repelled by every electron in the space between these electrodes. When the voltage across the tube is low, this *space charge*, as it is called, may thus actually check the rate of emission of the electrons. As the voltage is raised the electron flow increases also, the checking action due to the space charge repulsion becoming less and less until the voltage becomes high enough to draw off *all* of the electrons which the filament is emitting at the temperature at which it is operated. Thereafter a still further increase in the voltage has little effect upon the current flow.

Since the electrons move slowly on first leaving the filament they are much closer together there than some distance away from it towards the anode where they have accumulated considerable velocity. Thus the negative space charge is particularly dense near the filament where it can be most effective in checking electron flow. The presence of but few positive ions has a very great neutralizing action on electron space charge, since the relatively massive positive ions move much slower than electrons and remain between the electrodes much longer. One positive ion per second may thus neutralize the space charge effect of more than 100 electrons per second.

781. The Three Electrode Vacuum Tube. If a third electrode

FIG. 457. Space charge limited current
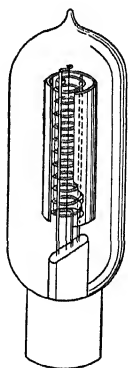
FIG. 458. Three electrode vacuum tube

in the form of a grid (Fig. 458) is placed between the filament and plate, as the anode is called, of a vacuum tube operating at such a voltage that its electron current is limited by the negative space charge, it is found that the slightest rise and fall of the potential of this grid above and below the filament produces a decided rise and fall of the current from filament to plate, called the plate current. This is not true, however, for a voltage between filament and plate which is so high that the full electron current is permitted to flow. In the latter case a slight fluctuation of the potential of the grid has no appreciable effect upon the electron current. Three electrode vacuum tubes are devices by which a slight voltage fluctuation, through the space charge fluctuation it produces, may be transformed into an exactly corresponding fluctuating electron current. The voltage fluctuation is applied between grid and filament, the latter usually being at ground potential. These tubes are therefore always operated at voltages low enough so that the plate current is limited by space charge. The best operating voltage depends upon the tube design.

782. X-rays. A very remarkable kind of radiation emanating from those parts of a cathode ray tube which are bombarded by the electron stream was discovered by the German physicist,
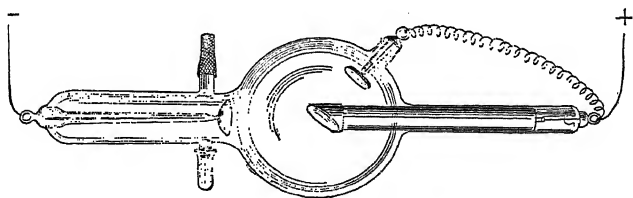
FIG. 459. Röntgen-ray tube

Röntgen, in 1895. This radiation, called Röntgen rays or X-rays is strongly sent out from an oblique plate of tungsten or platinum on which the cathode rays are converged, as shown in figures 459 and 460.

X-rays are detected by photography, as they act powerfully on an ordinary dry plate, and also by their power to excite fluorescence. They are not deflected by a magnet. They are reflected to some extent, however, and are slightly refracted by prisms

and lenses, although this latter is so slight that it was but recently detected.   They have a remarkable power of penetrating sub-stances of small density such as wood, pasteboard or flesh, which are opaque to light.   While they penetrate metals to a less extent, especially those of large density like lead and platinum, they nevertheless are being used for the detection of gas bubbles and other flaws on the interior of metal castings.

A fluorescent screen made of pasteboard covered with fine crystals of barium platino cyanide or calcium tungstate will glow brightly if brought in front of an X-ray tube in a darkened room. If the hand is interposed between the tube and the screen, the flesh, being most easily penetrated by the rays, will show but
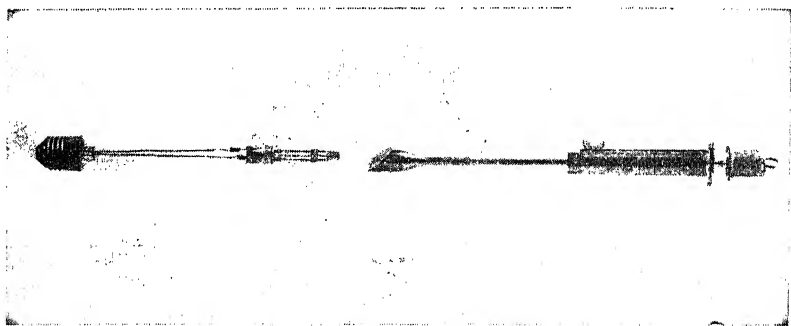
FIG. 460.   Coolidge X-ray Tube

faintly, while the shadow of the bones is strongly marked.   If a photographic plate enclosed in the usual plate holder with slides of hard rubber, wood or pasteboard is substituted for the fluorescent screen, a photograph is obtained on development such as shown in figure 461.

A mass of gas is ionized and made conducting by exposing it to X-rays.   When a charged electroscope is exposed to X-rays it quickly loses its charge.

783. Coolidge X-ray Tube.   The older forms of X-ray tubes as shown in figure 459 employed a cold cathode and contained a small amount of gas necessary to make them conducting as de-scribed in § 773.   A small variation in the amount of gas in the tube or the presence of small amounts of gases which may come from the electrodes or the glass have a great effect upon the

ionization produced and change the electrical conductivity of the tube so that its operation is difficult to control. Such a tube cannot be operated at very high voltages without excessive electrical discharges passing through it.

This difficulty was overcome by W. D. Coolidge who used a highly evacuated tube with a hot cathode to furnish the necessary electron supply for the cathode ray beam (Fig. 460). The voltage which may be applied to such a tube is limited only by its strength to withstand rupture or spark-over. A large Coolidge X-ray tube may be operated at over 200,000 volts. Further-
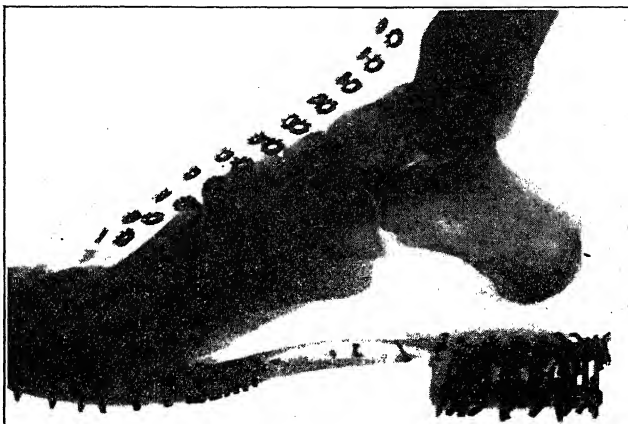
FIG. 461. Radiograph of foot in shoe

more, accurate control of electron discharge is possible by regulating the filament temperature. The high voltage used drives the electrons against the anode with great velocity and produces X-rays of very great *hardness*, or penetrating power.

X-ray tubes of this type must be carefully screened with sheet lead to intercept stray X-rays which may otherwise produce burns of a very serious nature upon the flesh of the operator.

784. The Nature of X-rays. It has been proved that X-rays are a wave phenomenon, that is, they are transmitted through space as waves of exactly the same nature as waves of light (§ 1006) but they are waves of extraordinarily short wave length, far too short to be perceived by the eye as light, except through

secondary effects as described in § 782. It will be seen later that X-rays, light waves and radio waves are all of the same nature; they are electromagnetic waves. X-rays are more fully discussed under light (§§ 977–980).

785. Positive Rays. In contrast to cathode rays, which are a beam of electrons, beams of relatively massive positive ions are also observed, when the cathode in a highly exhausted vacuum tube is pierced with holes. These *positive* rays as they are called may be seen as streams of light coming through the holes and out back of the cathode, and thus are directed away from the positive electrode in the opposite direction to the cathode rays. They were discovered in 1886 by Goldstein and originally called canal rays because of the small openings or canals through which they passed. The positive ions which make up these rays are produced through ionization by collision from the small amount of gas remaining in the tube. The exhaustion must be high enough, however, so the positive rays are not too much interfered with by collisions with gas molecules. It must be higher than for cathode rays because of the relatively large size and slow speed of the positive ions. The positive ion of hydrogen has a mass of nearly 2000 times that of the electron, and positive ions of other elements are still more massive. Therefore the velocity of positive rays is very much less than that of the cathode ray particles. These canal rays are found to consist of positively charged particles by experimental tests similar to those which established the identity of cathode rays (§ 776).

If a single beam of the positive rays is projected through a strong magnetic field the beam can be made to spread out into separate beams, depending upon the number of kinds of positive ions present, since the lighter ions are more deflected than the heavier ones by the magnetic field. By this type of *positive ray analysis*, J. J. Thomson, F. W. Aston and A. J. Dempster have proved the existence of and measured the atomic weights of *isotopes* of elements. These are discussed in § 806.

RADIOACTIVITY

786. Radioactivity. In 1896 the French physicist Becquerel discovered that minerals containing uranium gave out a radiation which resembled X-rays in acting on a photographic

plate through an envelope of black paper. He also showed that the uranium or Becquerel rays, as they were called, had the power to discharge electrified bodies, so that by using a sensitive electroscope they could be easily and accurately detected and their intensity measured. The photographic action was very slow, an exposure of several days being required to produce a distinct impression.

Following out this discovery, Mme Curie made a systematic search for other active substances and found that thorium possessed a similar power of *radioactivity*, as it now came to be called, a discovery which was also independently made by Schmidt.

It also appeared from these investigations that radioactivity is an *atomic* phenomenon. For the radioactivity of any given compound of uranium was found to be simply proportional to the amount of uranium in the substance, and in no way dependent on its physical or chemical condition.

But *uraninite*, an oxide of uranium called pitchblende, was found to be several times more active than could·be accounted for by the amount of uranium which it contained, and hence Mme Curie concluded that it must contain some unknown and highly radioactive substance, and resolutely set out to isolate it. As a result of the laborious treatment of several tons of pitchblende and uranium residues, a few hundredths of a gram of a new and astonishingly active element were obtained (in 1898) to which the name *radium* was given. This substance is obtained usually as a chloride or bromide, and is estimated to have in the pure state about two million times the activity of uranium.

Some radium compounds glow with a faint luminosity in the dark, though pure radium bromide is only feebly luminous. Crookes showed that if a minute particle of radium bromide is supported about a millimeter in front of a surface coated with phosphorescent zinc sulphide, the latter lights up with flashes of light, probably due to its bombardment by alpha particles from the radium. The scintillations which make up the glow on luminous watch hands are produced by the emission of alpha particles from a radioactive material used in the luminous coating.

787. Complex Character of the Radiation. The radiation from uranium, thorium, or radium, is found to contain three distinct kinds of rays known as alpha, beta, and gamma rays.

These different rays are distinguished from each other in two ways: by their penetrating power and by their deflection in a magnetic or electric field.

The *alpha* rays are completely stopped by a few centimeters of air or a layer of aluminum foil 0.05 mm. in thickness. They consist of positively charged particles having a mass about four times that of the hydrogen atom and are projected with a velocity of about 20,000 miles per second.

The *beta* rays have about 100 times the penetrating power of the alpha rays. They consist of negatively charged particles having only about $\frac{1}{1840}$ the mass of the hydrogen atom and seem to be of exactly the same nature as the cathode rays in a vacuum tube except that the velocity of projection of the particles in the *beta* rays is much greater than is usual in cathode rays, and in case of radium is found to be from 0.3 to 0.9 the velocity of light.

The *gamma* rays are the most penetrating of all; these rays from 30 mg. of radium bromide having been detected by their effect on an electroscope after passing through 30 cm. of iron. They are not deviated in a magnetic or electric field and are of the same nature as the Röntgen rays from a "hard" X-ray tube. They probably originate in the collisions of *beta* particles with the substance itself in its interior, just as X-rays arise from the impacts of cathode-ray particles against some obstacle.

788. Ionizing Power of the Rays. The power to discharge electrified bodies is possessed by all three kinds of rays, and is explained by their ionizing effect (§ 768) upon the gas through which they pass. The *alpha* particles with their comparatively great momentum have the greatest ionizing power and as they rush through a gas knock out electrons from an immense number of atoms along their paths. The *beta* particles or electrons which are ejected also ionize some of the atoms which they strike, but even though their velocity is very great, because of their smaller mass they produce less ionization than the *alpha* particles. The *gamma* rays or X-rays also have the power to make a gas conducting through the ionization they produce.

789. Paths of Alpha and Beta Particles. The ionizing effect of alpha and beta particles upon a gas has been demonstrated

in a very remarkable series of photographs obtained by C. T. R. Wilson, two of which are reproduced in figures 462 and 463. The method which he employed makes use of the fact that when a gas saturated with moisture is suddenly expanded a cloud of condensed vapor forms. Wilson found that by regulating the exact degree of expansion the condensation might be caused to take place only on atoms of gas that were ionized, so that, since an alpha particle in shooting through a gas produces a great number of ions along its path, the sudden expansion of the gas just after the particle has passed causes the condensation of microscopic drops on these ions and if these drops are instantaneously illuminated by an electric spark, before they have time to scatter or dissipate, they may be photographed and the path
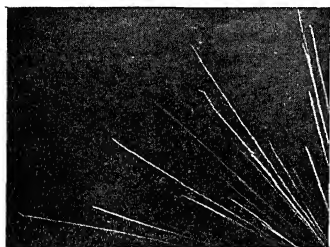
FIG. 462

FIG. 463

of the particle thus outlined. Figure 462 shows the intense ionization caused by the motion of an *alpha* particle from radium, something like 30,000 ions per centimeter being produced along its path. As its velocity grows less its path is more affected by the atoms through which it strikes, as shown by the sudden changes in direction toward the end. Finally, its velocity becomes too small to ionize the gas and its path becomes invisible.

Figure 463 shows the paths of *beta* particles which are set free from atoms of gas by a narrow pencil of X-rays. Each beta particle produces but few ions along its course compared with the number produced by an alpha particle, and their paths are very crooked and irregular because the momentum of one of these particles is so small that its direction of motion is greatly changed in passing through atoms of gas.

790. Energy and Heating Effect of Rays. Although the velocity of the *alpha* particles is less than that of the *beta* particles, yet in consequence of their greater mass *the energy of motion of each alpha particle is something like* 120 *times that of a beta particle*, consequently the heating and ionizing effect of the rays is mainly due to the *alpha* particles.

It was found by Curie and Laborde, in 1903, that radium was a constant source of heat, one gram of radium bromide giving out more than 100 gram-calories of heat per hour, or more than enough to melt its own weight of ice in an hour. This extraordinary development of energy is believed to be mainly due to the escaping alpha particles which are absorbed within the mass of radium itself and in the enclosing vessel, their energy of motion being transformed into heat. Originally the energy must have existed in the radium atoms before the alpha particles were given off.

791. Radioactive Transformations. Evidently an atom of radium cannot give out alpha and beta particles with a corre-
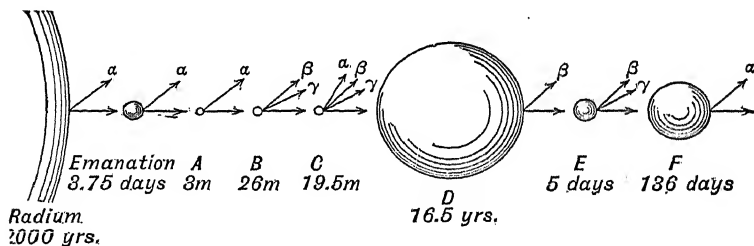
Fig. 464.   Diagram of radioactive transformations

sponding expenditure of energy and remain the same as before. The atoms of the radioactive substances have such slight stability that at every instant some of them are reaching a condition of instability and breaking up or exploding with the expulsion of alpha or beta particles and coming to a new state of equilibrium. In case of radium, about half the particles in a given mass will have broken up in this way in the course of 2000 years. When such a change has taken place, and an alpha particle has been expelled, the new state of equilibrium is often even less stable than the original one, and consequently another alpha particle is soon

expelled and another state of equilibrium reached which itself may also be short-lived, and so there takes place a series of changes in the course of which the various kinds of rays are given off, and finally a condition of such stability is reached that no further change is detected. This disintegration theory was proposed by Rutherford and Soddy in 1902.

The diagram indicates the series of changes of the radium atom as followed out by Rutherford, and under each step is given the time in which that product is half transformed. Thus, while radium itself is half transformed in 2000 years, the instability of the emanation is so great that any given portion of it will be half transformed in 3.75 days; while radium A is the most unstable of all, its period of half transformation being only 3 minutes.

The sizes of the spheres in the diagram represent the relative amounts of the various products that are present in a mass of radium which is undergoing change, and has come to such a state of equilibrium that each product is being produced from the one preceding it in the series, just as rapidly as it is being transformed into the next succeeding one, and is therefore neither increasing nor decreasing in amount.

Since radium is being constantly though slowly transformed, it must disappear from the earth in the course of a few thousand years unless it is being produced in some way. It is now known that radium is a disintegration product of uranium, for uranium has the greater atomic weight, they are always found together, and it has been shown that the amount of radium in any radioactive mineral always bears a constant ratio to the amount of uranium which it contains, viz., about 1 to 2,630,000.

The transformation of uranium itself is so slow that it will require, according to Rutherford, a period of at least ten thousand million years for any large fraction of it to be transformed.

Uranium goes through five transformations before radium is finally reached. The last two products before radium, called Uranium 2 and Ionium, have each a life of more than 100 times that of radium.

792. Helium. In 1868 Lockyer proposed the name *helium* for an unknown substance existing in the sun and causing a line in the solar spectrum which could not be obtained from any

known terrestrial substance. Nearly thirty years later Ramsay, an English chemist, identified it with a gas obtained from some radioactive minerals. Its occurrence in these minerals suggested to Rutherford and Soddy that it might be a product of the disintegration of radium. Later Ramsay and Soddy collected in a tube some of the emanation from radium, and though at first there was no indication of the presence of helium, after five days its complete spectrum was obtained. Helium is a gas having twice the density of hydrogen, and it is found that the alpha particles given off in the transformations of radium form the atoms of helium. This result is remarkable as the first case in which one stable element has been derived from another.

793. Final Product of Radium. If the atomic weight of the alpha particle is 4 (the same as helium), then, since the atomic weight of uranium is 238.5, the loss of three alpha particles would bring it down to 226.5, which is almost the same as 225, the observed atomic weight of radium. Then, as five alpha particles are given off from radium in the series of changes indicated above (§ 791), the atomic weight of the final product would be expected to be 206.5, which is in close agreement with 206.9, the atomic weight of lead. The suggestion that lead may be the final product of the radium changes is due to Boltwood and is supported by the fact that lead is always found associated with radioactive minerals which are rich in uranium.

It has been shown recently that lead from radioactive minerals has an atomic weight slightly less than that of *ordinary* lead. The two kinds of lead differ slightly in density, but in their properties including their spectra they are identical.

This difference in atomic weight is explained through the presence of several isotopes (See § 806) in ordinary lead the mixture of which gives it a somewhat greater atomic weight.

These discoveries that radium, helium, and lead also are products of the disintegration of uranium give new emphasis to a very old suggestion that the atoms of what we call the elements are a number of stable aggregates built up from some simple primordial atom. Whether the elements as we now know them are to be regarded as so many stages in the gradual disintegration of originally more complex atoms of high atomic weight or are various stable aggregates formed independently and perhaps si-

multaneously under physical conditions that we cannot now distinctly conceive, is a question that cannot now be answered, though the absence of any perceptible radioactivity in the case of most substances points toward the latter origin. These ideas are further developed in § 806.

794. Internal Energy of Atoms. That there exists an enormous amount of energy in the interior of atoms is evident from the heat which is given out by radium in its slow transformation. It has been calculated that one gram of radium gives out 133 gram-calories of heat per hour, but as the *average life* of a radium atom is 2440 years, it follows that the whole amount of heat emitted by one gram of radium in the process of transformation is 2,830,000,000 gram-calories or more than three hundred thousand times as much as is given out in the combustion of a gram of coal. And it must not be forgotten that *this enormous amount of energy can represent only a small fraction of the total internal energy of the atoms in a gram of radium*, being merely *that part which is given out when the atoms pass from one state of equilibrium to another*. This energy is detected in radioactive elements only in consequence of their disintegration, but there is no reason to suppose that the internal energy of these elements is of a different order of magnitude from that of the other elements.

ATOMIC STRUCTURE

795. All Atoms Contain Electrons. Electrons are associated with so many electrical, physical and chemical phenomena that they are without doubt a fundamental unit in the structure of all atoms of matter.

Cathode rays are always a stream of electrons, whatever metal the cathode which emits them is made of, or whatever be the residual gas in the tube. Charges of electricity produced by friction as in Millikan's oil drop experiments (§ 560) are found to be made up of grains of electricity identical with electrons. The *beta* particles emitted from radioactive substances, the negatively charged particles appearing during ionization (§ 768) or from metals acted upon by X-rays or by ultraviolet light are all found to have exactly the mass and charge of the electron. Since the mass of the electron is only about $\frac{1}{1840}$ of that of the hydrogen atom its charge is so small that it takes over 2000

million of them to make up the electrostatic unit as defined in § 540.

796. All Atoms Contain Positive Electricity; Protons. It was seen in the discussion in electrostatics that when electrons are removed from a conductor it is left positively charged. It is thus natural to suppose that atoms contain positive electricity as well as electrons. This positive charge, however, unlike electrons, has never been separated from the atom. All that can be said is that the positive hydrogen ion of mass about 1840 times that of the electron is the least massive separate positive charge known, its charge being equal but opposite to that of the electron. It has received the name *proton*. The modern theory of matter states that atoms of the elements are made up of nothing more than aggregates of discrete portions of positive and negative electricity. Mass inertia is the resistance offered by these electric charges to being accelerated. The reason why the mass of the positive elementary charge, or proton is greater than the mass of the negative charge, or electron, although the two charges are equal in amount, can be shown to arise from the much smaller volume occupied by the proton.

It is supposed that *in the neutral state every atom of matter is made up of a certain number of protons and an equal number of electrons*. When an atom loses an electron it becomes positive, when it gains an extra one it becomes negative.

797. The Structure of the Atoms. Every atom is believed to have a definite structure consisting of an extremely small *nucleus* in which is packed all the positive electricity and around which the negative electrons are arranged in some definite order but with wide spaces between them.

The *nuclear* theory of atomic structure is largely a consequence of some remarkable experiments by Sir Ernest Rutherford about the year 1912, in which *alpha* particles (§ 787) projected from radioactive substances were shot through thin foils of various metals. The number of alpha particles projected through a given small area of foil were actually counted from the scintillations produced by the alpha particles on a zinc sulphide screen. It was found that a certain small percentage of the alpha particles, instead of passing straight through the foil, were given sudden deflections so their directions on coming out made angles

with their directions on entrance. It is concluded that the alpha particles shoot *through the atoms* of the metal foil because these atoms are closely packed together; furthermore, that the atoms must be of very open structure and the alpha particle very small because every alpha particle driven through the metal foil must pass through thousands of the metal atoms. It is only very occasionally that an alpha particle is deflected in its course by some sort of encounter. Rutherford has shown that these results can be satisfactorily explained if each atom of the foil is assumed to have an extremely small positively charged nucleus in which practically its entire mass is concentrated, and also if the sudden deflections of the alpha particles on their near approach to these atomic nuclei are produced by an electric repulsion which follows the usual law (§ 539). Another requirement is that the alpha particles, too, be extremely small and of not very different size from the nuclei of the atoms of the foil, in fact the alpha particle is known to be identical with the nucleus of a helium atom (§ 792).

798. Size of Atomic Nuclei. From these results Rutherford actually calculated just how close the centers of the alpha particles approach to the centers of the atomic nuclei in passing through foils of different metals. He found that for gold foil this distance was about 3×10^{-13} cms. and that it did not vary much from this for other metals of very different atomic weight. The results seemed to show that the radii of the atomic nuclei of different elements do not vary far from 10^{-13} cms., those of high atomic weight having perhaps two or three times the diameter of those of low atomic weight. It has been found in different ways that *atomic* diameters are not far from 10^{-8} cms. (§ 281). Thus the astonishing result was obtained that the nucleus of an atom was about $\frac{1}{100,000}$ of the diameter of the atom, also that into this minute nucleus is packed all of the positive electricity and all of the mass of the atom except for that small part due to the electrons arranged outside the nucleus.

799. Size of the Electron. Electrons also are found to be extremely small compared with the size of the atom because they also can be shot through thin foils and through thousands of molecules of a gas without suffering encounters. Beta particles emitted from radioactive substances are identical with electrons.

The great distances through which alpha particles emitted from a radioactive substance will travel through gases at the atmospheric pressure without appreciable deflection shows that these particles must shoot straight through the molecules of the gas. Figure 462 shows the paths of such particles. For straight paths of these lengths, taking account of the diameter of the gas molecules, calculation shows that the alpha particles must *pass through several hundred thousand molecules* before their motion is arrested. Occasionally a sharp deflection is produced by some sort of impact, but such deflections are so rare compared with the number of molecules traversed that the masses of the alpha particles and those of the gas molecules must be concentrated in centers of extremely small size compared with the sizes of the atoms or molecules. Except for these small centers the molecules must be of very open structure. Since the alpha particle, now known to be the nucleus of a helium atom, has over 7000 times the mass of an electron, impacts with electrons would have little effect on its motion through the gas, and many electrons are dislodged from the gas molecules along the path of the alpha particle as negative ions (§ 768) without deflecting the alpha particle. But photographs of the paths of beta particles (which are the same as electrons) show that, although often deflected because of their small mass, probably due to the repulsions of other electrons in the molecules through which they pass, these beta particles may shoot through thousands of atoms and molecules without producing a single ion, or, in other words, without dislodging a single electron from a molecule or atom. Therefore these beta particles, or electrons, are also very small compared with the size of atoms and molecules.

800. Atomic Numbers. Another important result was obtained from the experiments on the scattering of alpha particles shot through metal foils. It was not only possible to calculate the approximate size of the atomic nuclei of the foil metal but to determine the amount of positive charge which they must carry. It was found that the number of elementary positive units of charge carried on each nucleus, the elementary positive charge being equal but opposite to that carried by the electron, was roughly *one-half the atomic weight of the elements*. This result is largely responsible for the suggestion in the year 1913 that when the elements are arranged in a series in the order of their increasing atomic weights beginning with 1 for hydrogen and running up to 92 for uranium, *the number corresponding to each element in this series is exactly equal to the number of elementary positive charges on its atomic nucleus*. The numbers assigned to the elements in such a series roughly correspond to one-half the atomic weights of the elements. For neutral atoms,

the number of *electrons* out of which the *outside structure* of each atom is built must also be equal to the number in this series corresponding to each atom because in a neutral atom the number of electrons and the number of elementary positive charges are equal. The number of an element in this series is called its *atomic number*. The probable truth of this idea was brought out by the remarkable results of Moseley's work on characteristic X-rays. When the anode in an X-ray tube, against which the cathode rays or electrons are driven by the applied voltage, is made of different metals it is found that the resulting X-rays emitted from each metal have certain special wave lengths which are called characteristic wave lengths. Moseley found that *the frequency of vibration of the characteristic X-rays of any element is proportional to the square of its atomic number*.

The *atomic number* of an element is even more fundamental than its atomic weight, because this number is exactly equal to the number of positive charges on the nucleus, and it is this positive charge which determines the number and the arrangement of the outside electrons which determine largely its physical and chemical properties.

801. **Bohr Theory of Arrangement of Electrons Around Atomic Nuclei.** Various theories of the probable arrangement and the motions of the electrons about the positively charged atomic nuclei have been proposed both by physicists and chemists. A very remarkable theory is one proposed by the Danish physicist, Bohr, by which the electrons revolve about the central nucleus in certain particular orbits which may be approximately circular or elliptical in form. The hydrogen atom has 1 electron revolving about its nucleus, the helium atom 2, the lithium atom 3, etc., up to uranium, which has 92 electrons revolving about its nucleus. In each case the number of electrons corresponds to the number of positive charges on the nucleus or to the atomic number. Each electron may revolve in any one of a certain particular set of orbits. The speed of revolution of an electron in each orbit is determined by the centripetal force exerted on the electron by the nucleus. For inner orbits near the nucleus where the attraction is strong the speed of revolution of the electrons must be very great. This theory was built up in an attempt to account for the lines in the spectrum of

luminous gases (§ 936). It is a development from Planck's quantum theory (§ 484).

According to Bohr's theory these radiations result when the electrons *fall from an outer to an inner orbit.* In a very hot luminous gas, for instance, the agitation of the molecules and electrons results in electrons continually being knocked from inner to outer orbits, but the electrons soon drop back to their inner orbits giving out again in the form of radiation the energy received by them. The highest frequency possible in a given atom is that resulting from electrons falling from a point outside of the atom to the very innermost orbit. Only the most powerful X-rays and high velocity cathode rays can knock out the electrons from the innermost orbits of the heavier atoms, which must take place before this high frequency radiation can occur. Much lower frequencies in the visible light region may result due to the dropping back of the outermost or *valence* electrons into their orbits again as in the case of recombinations of the electrons with positive ions after ionization such as may take place in a Geissler tube. Each revolving electron has a certain amount of kinetic energy due to its velocity of motion in its orbit around the atomic nucleus. It also has a certain amount of potential energy due to its distance from the attracting nucleus. The sum of the kinetic and potential energies of an electron is a constant for a given orbit. In fact, the most fundamental method of designating an electron orbit is by the constant amount of energy an electron has in that orbit. The larger the orbit the greater is the electron energy associated with it. Since each separate orbit in which an electron may revolve corresponds to a certain amount of energy, these orbits may be represented by a series of separate horizontal lines of heights above a zero line proportional to the orbit energies. These heights are known as *energy levels,* and form a characteristic series for each atom. The lower levels correspond to inner orbits and the higher levels correspond to outer orbits, and a still higher series of levels corresponds to orbits which displaced electrons may occupy temporarily before dropping back into permanent orbits again. It is a remarkable fact that there is a close correspondence between energy level series and the spectrum line series (§§ 936–937) for various atoms. The explanation of spectral series in terms of energy levels of electron orbits is one of the great contributions of the Bohr theory.

802. Bohr Theory and the Quantum Theory. According to the quantum theory (§ 484), radiation energy is given out in separate minute units called quanta. Bohr assumes that when an electron drops from an outer to an inner orbit, or from a higher to a lower *energy level,* just one quantum of energy is given out whose magnitude is exactly measured by the decrease in the energy of the electron due to this drop. This means that the larger the energy drop of the electron the larger

the energy of the quantum which must be radiated. The quantum theory takes care of this because it supposes that the quanta may be radiated in different magnitudes which are proportional to the radiation frequency, the large quanta corresponding to the high frequency waves of short wave length, and the small quanta corresponding to the lower frequency waves of longer wave length. For instance, it was mentioned in the previous paragraph that X-rays, which are very high frequency radiations, are produced when electrons displaced from inner orbits drop back again and lower frequency radiations of visible light are produced when electrons displaced from outside orbits drop back again to those orbits. Each displaced electron gives out one quantum during its drop, but the large energy drop of the electron associated with X-ray emission produces a quantum which may contain thousands of times as much energy as the small quantum radiated when an electron makes a small energy drop.

803. The Lewis-Langmuir Theory of Arrangement of Electrons Around Atomic Nuclei. Another theory of atomic structure which has attracted wide attention is the Lewis-Langmuir theory first proposed by G. N. Lewis and developed by I. Langmuir. This theory also is built up on the fundamental idea of a positive nucleus, the number of positive electric units on the nucleus and the number of electrons in the outside structure being each equal to the atomic number of the element. It differs from that of Bohr in that it is intended primarily to account for chemical and physical properties of substances, and makes no attempt to account for atomic radiation phenomena. The electrons are considered to be arranged around the nucleus in concentric spherical shells. The innermost shell can have no more than 2 electrons in it, the second can have no more than 8. Hydrogen has only 1 electron in the inner shell; helium has 2 in the inner shell, one on each side of the positive nucleus, lithium has 2 in the inner shell, but its third electron must be in the second shell because the first shell can contain no more than 2; the gas neon of atomic number 10 has 10 electrons, 2 in the inner shell and 8 in the second shell. For elements of higher atomic number than 10 the electrons occupy more than 2 shells; in fact, elements of high atomic number are believed to have their electrons arranged about their nuclei in several shells.

The interest of this theory lies in the fact that by taking account of the various degrees of stability for the electron arrangement supposed for each atom, various chemical and physical properties of the elements can be accounted for and even predicted. Furthermore, this theory gives the best explanation of chemical valence thus far obtained, and gives a new meaning to the periodic table arrangement of elements.

Let us consider a few simple cases. Take neon first, of atomic number 10. Its nucleus has 10 positive charges on it, so that the first 2 shells are just filled, there being 2 electrons in the inner shell, 1 on each side of the nucleus, and 8 in the second shell, which may be assumed to be equally spaced from each other as though on the 8 corners of a cube. According to the Lewis-Langmuir theory, this is an arrangement of great stability, an electron being removed only with the greatest difficulty. A general law of this theory is that *electrons have a strong tendency to arrange themselves in complete shells.* Thus, for an atom like that of neon which has 2 complete shells, the electrons are held with great stability. When the electrons are held very powerfully practically no field of force is left outside of the atom. These atoms cannot then become attached to other atoms, resulting in the formation of chemical compounds. We should therefore expect neon to be a chemically inert gas and a monatomic gas because the atoms will not even cling together in pairs. This is what is actually true of neon.

We will next take sodium of atomic number 11. This is just 1 *greater* than that of neon so the neutral sodium atom has the first 2 shells full of electrons like neon but has 1 extra electron outside in the third shell. We should expect that this outside electron could be easily detached because this leaves the remaining 10 electrons arranged in 2 complete shells like those of the very stable neon atom. This is what actually happens in the case of sodium vapor. It is easily ionized, or its atom easily loses its extra electron, leaving it as an ion with 1 extra positive charge equal but opposite to that of an electron. We have also accounted for the fact that sodium is strongly electropositive, tending to form positive rather than negative ions.

Now let us consider fluorine of atomic number 9. This is just 1 *less* than that of neon so the neutral fluorine atom has to have 1 *more* electron in order to fill the first 2 shells giving the stable neon arrangement of 10 electrons. We should thus expect fluorine to have a strong tendency to pick up an extra electron to complete its second shell. But this would give an extra negative charge to the fluorine atom making it a negative ion. This is according to the facts, as it is well known that fluorine is strongly electronegative.

Fluorine is known to form molecules of considerable stability and therefore exists at ordinary temperatures as a gas, that is, its molecules do not tend to cling together to form a solid. This illustrates another important law of the Lewis-Langmuir theory. According to this law *pairs of electrons*

may be held in common by 2 *atoms* when such a combination completes the outside shells of the 2 atoms. Thus the neutral fluorine atom whose outside shell lacks 1 electron, tends to form neutral molecules of 2 atoms each in which the 2 separate atoms are held together because 2 *of the electrons are held by both atoms at the same time.* This gives an outside shell of 8 electrons for each atom, corresponding to the 8 corners of a cube but 2 electrons are held by both atoms at the same time corresponding to 2 cubes clinging together along 1 edge bringing 2 corners of each into coincidence. When 2 atoms are held together in this way by 2 electrons held in common we have an illustration of a *chemical valence bond.* Since the fluorine molecule is neutral and the electron arrangements are stable because the outside shell for each atom is complete there is only a small outside field of force left so the molecules remain separate from each other and the fluorine remains in gaseous form.

804. Limitations of the Theories. That neither of these theories is complete in itself is recognized by both physicists and chemists, but that each contains some truth can hardly be doubted. It can be said that the theories are not antagonistic or incompatible with each other as was at first thought before Bohr's theory reached its present state of development. The formation of chemical compounds in the case of the Bohr theory is conceivable even though the electron orbits of the different atoms have to intersect when the atoms combine with each other, to form molecules. It must be recalled here that the electrons are probably not more than $\frac{1}{100,000}$ the diameter of the atom. *If the atom were* 100 *yards in diameter, its nucleus and also the electrons arranged about it would be about the size of pin heads,* so electrons are not likely to interfere with each other when their orbits intersect.

805. Only Outermost Electrons of the Atom Take Part in Physical and Chemical Phenomena. According to these ideas of atomic structure, it is only the very outside electrons or valence electrons of the atom which take part in ordinary physical and chemical phenomena because these are the electrons most easily detached. In the case of metals of high electrical conductivity, certain of the outside electrons (called *free electrons*) are so loosely held to the atoms that they are passed on from one atom to the next by the very slightest potential difference, the resulting electric current consisting simply of a stream of electrons passing from atom to atom. It is an interesting and significant fact that electrical conductivity and heat conductivity closely correspond. Silver is the best conductor known for either electricity or heat. Heat conductivity is therefore believed to be a phenomenon primarily due to the *free electrons* passing on their heat vibrations from

one to another by their impacts. Electrons emitted from hot bodies (§ 770), those emitted from substances acted upon by ultraviolet light or those emitted in the form of cathode rays are all loosely held electrons.

Electrical insulators, which are also poor conductors of heat, have the outside electrons held firmly so they are not easily passed from atom to atom, but will merely spring forward slightly with an applied difference of potential and spring back again upon its removal without leaving the atoms (§ 575).

Atoms whose outside electrons are very stable and strongly held to the nucleus like those of the gases helium, neon and argon have practically no outside field of force left so they exist as separate atoms, only assuming a liquid and solid form at the lowest temperatures, in which condition their heat vibrations which tend to hold them apart are very slight. In the case of the metals which melt only at high temperatures, the atoms, although electrically neutral, have a structure such that the outer electrons are loosely held and a considerable outside field of force is left so that the atoms attract each other strongly and tend to build up solid crystalline forms at ordinary temperatures. When the atomic vibrations due to high temperature are violent enough, however, even these strong attractions are overcome, and the metal can exist only as a vapor or gas.

The formation of chemical compounds, some very stable and others much less so, and even surface tension phenomena (§ 265) and cohesion and adhesion (§ 236) are also believed to result from the same sort of electrical attractions between the atoms and molecules which depend upon their out-side electron arrangements.

All of these phenomena whether chemical or physical are simply different manifestations of electrical attractions between atoms and molecules, the attractions being governed by laws determined by the external structures of the atoms and molecules. These phenomena are therefore all fundamentally electrical in nature.

THE NUCLEUS OF THE ATOM

806. Isotopes. Positive rays or canal rays (§ 785) result from the ionization of atoms or molecules (§ 768) by which one or several of the outer loosely held electrons are knocked off, leaving the rest of the atom or molecule with one or several unneutralized elementary positive charges. These *positive rays* thus contain the nucleus and most of the electrons of an ionized atom, and therefore practically its entire mass, as it will be recalled that the mass of an electron is only about $\frac{1}{1840}$ of that of the hydrogen atom.

In studies of the deflections of these *positive rays* produced by projecting them through magnetic fields (§ 776), Sir J. J. Thomson at the Cavendish Laboratory in Cambridge, England, made the notable discovery that the gas neon may exist in two forms with identical chemical properties but with two atomic weights of exactly 20 and 22 as near as could be determined, the lighter ions being deflected just a little more than the heavier ones in passing through the magnetic field.

Since this discovery the same thing has been found to be true of many other elements chiefly as the result of the work of F. W. Aston, also at the Cavendish Laboratory, who used an ingenious modification of J. J. Thomson's method, this new method making possible much more accurate analyses. Lithium is a mixture of two types of atoms of atomic weights of exactly 7 and 6. Magnesium is a mixture of atoms of atomic weights 24, 25 and 26, the gas argon is a mixture of atoms of atomic weights of exactly 40 and 36, and many other similar cases have been discovered. Such different forms of the atoms of a single element which can only be distinguished from each other by their difference in atomic weights are called *isotopes*. Since the physical and chemical properties of the isotopes of a given element are identical except for their differences in atomic weights they had never previously been separated by ordinary methods of analysis.

But the most remarkable fact of all which has appeared is that, except for hydrogen, no isotopes have ever been found whose atomic weights are not almost exactly whole numbers. It is believed that the reason the atomic weights of the elements as found by older methods do not come out as whole numbers in many cases is that each atomic weight represents an average for two or more isotopes of the element mixed in different proportions. The separate isotopes, however, all have atomic weights which are whole numbers.

807. Atomic Nuclei. That the number of elementary positive charges on the nucleus of the atom is probably more fundamental than the atomic weight in determining its properties has already been mentioned (§ 800). The discovery of isotopes lends interesting support to this view. For instance, the atomic nucleus of the gas neon has 10 elementary positive charges on it. This means that the neutral atom must have 10 electrons arranged in the outside structure around the nucleus. It is these 10 electrons held by 10 positive charges on the atomic nucleus which determine practically all the physical and chemical properties of neon, and the mere fact that the atomic *weight* may be 20 or 22 as in the case of two isotopes of neon is a subordinate matter. The nucleus of the atom, it will be recalled, although so minute, contains practically its entire mass, and therefore isotopes probably represent differences in structure of the atomic nuclei, which have no effect on the physical and chemical properties of the atom so long as the number of elementary positive charges on the nucleus which determine the atomic number remains the same.

Since isotopes have atomic weights which are all whole numbers (except hydrogen which is 1.008) the possibility of the atomic nucleus itself being built up of separate units at once appeared. In confirmation of this idea, Rutherford has found that when swift alpha particles from a radioactive substance are projected through air or nitrogen, particles appear having the atomic weight and the single positive charge of the hydrogen atomic nucleus, which travel such distances through the gas that it is impossible to believe that their motion is not due to something more than a collision with an alpha particle. It is believed that these hydrogen nuclei are actually expelled with great violence from the nucleus of the nitrogen atom whose

structure is broken down by a blow from an alpha particle, which may be traveling with a velocity of over 20,000 miles a second (§ 787). Boron, fluorine, sodium and several other elements are also found to eject hydrogen nuclei from their atomic nuclei when bombarded by alpha particles.

Radioactivity also furnishes evidence as to the probable complexity of the atomic nucleus. The alpha particles and beta particles which are nuclei of helium atoms and electrons (§ 792) are believed to come from the nucleus of the atom of the radioactive element, because the helium atomic nuclei and electrons are emitted in exactly the same way, no matter what be the temperature of the radioactive substance or how it is chemically combined. Radioactive changes are probably the gradual breaking down of atomic *nuclei* from complex to simpler forms (§ 793).

808. Protons. These discoveries taken in connection with radioactive phenomena suggest that perhaps the *nucleus of the hydrogen atom* may be taken as the ultimate positive unit of electricity. It carries a positive charge which has never been separated from it which is exactly equal but opposite to that carried by the electron.

The name *proton* has been suggested for this ultimate positive unit of electricity, identical with the hydrogen nucleus and therefore with a mass over 1800 times that of the electron. Practically all of the mass of the atom as well as all of its positive charge is packed into its nucleus. Take the case of neon, for instance, of atomic number 10 and isotopes of atomic weight 20 and 22. The first isotope has 20 protons packed into its nucleus to account for its mass and it also has 10 electrons packed in which neutralize the charges of 10 of the 20 protons, leaving the nucleus with a charge of 10 positive units corresponding to the atomic number. The second isotope contains 22 protons and 12 electrons packed into its nucleus to account for its mass and nuclear charge.

It is believed that *the atomic weight of every isotope is exactly equal to the total number of protons in its nucleus,* and the *atomic number is the number of positive units of electricity remaining on the nucleus after the charges of about half the total number of protons have been neutralized by electrons.*

These electrons which are packed into the nucleus with the protons have nothing to do with those electrons which go to make up the structure outside of the nucleus. The external electrons are arranged far outside of the nucleus with wide spaces between them. It is with the arrangements of these external electrons only that the Bohr theory and the Lewis-Langmuir theories are concerned. The study of the structure of atomic *nuclei* covers quite a different field.

809. Loss of an Electron from Atomic Nucleus Raises the Atomic Number One, Loss of a Proton Lowers It One. The atomic number or the number of positive charges on the nucleus of the atom of an element, it will be recalled, is what determines all the physical and chemical properties by which it is identified. Therefore, the loss of an electron or of a proton from the nucleus should transform the element to one of next higher atomic number or to one of next lower atomic number, because the loss of a nuclear electron leaves one more positive charge on the nucleus, and the

loss of a proton leaves one fewer positive charge on the nucleus. This is just what is found to happen in radioactive changes, which are known to be the result of the breaking down of atomic *nuclei* and have nothing to do with the external electron structure. When a beta particle or an electron is ejected, the atomic number of the atom whose nucleus gives it out goes up 1, and the atomic weight remains practically the same because the mass of the electron is so small. When an alpha particle is ejected, however, the atomic number goes down 2 because the alpha particle is the nucleus of the helium atom of nuclear charge 2. The mass of the helium atom is 4, however, so the atomic weight at the same time decreases by 4. If an alpha and a beta particle are given out from an atomic nucleus at the same time, the atomic number decreases by only 1, but the atomic weight by 4 as in the previous case. The final product of the breaking down of radium (§ 793) is an isotope of lead, as it has exactly the same nuclear charge as the lead atom has, though a somewhat different atomic weight. The helium atom is considered to have four protons and two electrons packed into its nucleus to account for its mass and nuclear charge. It should be noted that in all known radioactive changes, the protons are not given off as individuals, but in pairs in the form of alpha particles.

810. Stability of Atomic Nuclei. The only way the structure of the nucleus of the atom has ever been known to be broken down by an external agency is by Rutherford's method of alpha particle bombardment. Radioactivity represents a breaking down of atomic nuclei, but this is an automatic process which it has never been possible to control.

ELECTRIC OSCILLATIONS AND WAVES

811. Oscillatory Discharge of a Leyden Jar. It has been already stated (§ 598) that when the resistance of the discharge circuit is sufficiently small the discharge of a Leyden jar is oscillatory. This was discovered by the American physicist, Joseph Henry, who, as early as 1842, found that when a Leyden jar was discharged through a wire wound around a needle the latter was magnetized, but sometimes one end was made the north pole and sometimes the other, although the jar was always charged the same way. He believed that this was caused by the oscillation of the discharge current which kept reversing the magnetism of the needle back and forth until the current became too small to have a further effect.

Lord Kelvin, in 1855, quite unaware of Henry's discovery, showed by the principle of energy that the discharge must oscillate back and forth until all the original energy of charge is expended in sound, heat, light, and radiation, and that when the

resistance of the circuit is very small the period of oscillation is given by the formula

$$T = 2\pi\sqrt{LC} \, *$$

where L is the inductance of the circuit and C is the capacity of the jar (compare with vibration formulas of §§ 135 and 149). In case of an ordinary gallon jar discharged by a short discharging rod, the period of oscillation may be as small as two ten-millionths of a second, while Lodge, by using a battery of large capacity and discharging it through a very long circuit having large self-induction, was able to make the alternations so slow as to give out a distinct musical note. Feddersen, in 1859, first analyzed the spark by a rotating mirror, as already related (§ 598).

812. Electric Resonance. When a Leyden jar is discharged not only may there be oscillations in the discharge circuit itself, but in consequence of induction there are set up electric oscillations or surgings in neighboring conductors. In general these are but feeble, but if the free period of the surging happens to be the same as that of the oscillations in the discharge circuit, quite energetic surgings may result, just as a tuning-fork will excite strong vibrations in a resonator which is in tune with it. *The circuits are then said to be in resonance* (§ 134).

813. A Case of Electrical Resonance. The influence of electrical resonance is well shown in the following experiment due to Lodge. Two Leyden jars of nearly equal capacities are chosen. One which can be charged by an electrical machine or induction coil is provided with a short circuit of thick wire which is attached to the outer coating and terminates in a knob separated by a short spark gap from the knob of the jar. The second jar has a strip of tinfoil reaching from the inner coating over the edge and terminating in a point at e near the upper edge of the outer coating; its inner and outer coatings are connected by a

---

* This formula is exactly analogous to those for mechanical oscillations (§§ 135 and 149) where the inductance L corresponds to mass and the *elastance* $\dfrac{1}{C}$ corresponds to the restoring force per unit displacement. The inductance may be looked upon as electrical inertia which carries the current along during the oscillation, and the elastance as an elastic restoring force acting upon the charge of electricity causing it to recoil.

wire circuit, part of which, marked AB in the figure, can be slid along changing the length of the path. When the two jars are placed, say, a foot apart with the two circuits parallel, a position for the slider AB may be found by trial, such that whenever the first jar discharges across between the knobs, a spark leaps the gap between the tinfoil strip and the outer coating of the second jar. If the slider is moved a short distance away from this position in either direction, the sparks at e cease. Lodge calls the sparks at e the "slopping over" of the powerful surgings due to the two circuits being in resonance.
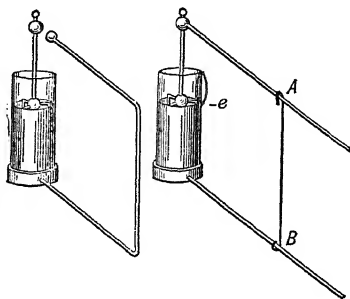
FIG. 465. Sir Oliver Lodge's resonance experiment

814. Electric Waves in Wires. When one end of a long straight wire is given a charge or touched to a battery pole, a wave of electric pressure or potential runs along the wire with a velocity which depends on the insulating medium immediately surrounding the wire. *In case of a straight bare wire in air the wave has the velocity of light;* but when the wire is coiled, forming a closely wound helix, the wave travels much slower on account of the greater inductance of the coil.

On reaching the end of the wire the wave is reflected back, just as a sound wave is reflected at the end of a stopped organ pipe. If, instead of a single impulse, a series of alternate positive and negative charges are given to the end of the wire in exactly the right frequency, it may be set in strong electrical resonance just as a stopped pipe vibrates powerfully when a tuning-fork of the proper frequency is sounded at its mouth. Resonance will occur when the period of the electrical impulses is four times as long as it takes a wave to run the length of the wire, exactly as in case of a stopped organ pipe.

The resonance of waves in wires may be beautifully shown by the following experiment due to the German electrician, Seibt:

A large Leyden jar has its coatings connected by a circuit having a spark gap at S with zinc knobs. By moving the slider L nearer to the jar or farther away, the length and self-induction

of the discharge circuit may be varied and consequently the period of the oscillatory discharge can be adjusted.

Two long helical coils of wire A and B are mounted on insulating stands. They are both connected at the bottom to one of the coatings of the Leyden jar while each terminates above in a point. One helix is wound with a much greater length of wire than the other.

If by means of a powerful induction coil the Leyden jar is caused to discharge across the gap S, each discharge will be oscillatory and consequently a series of impulses is communicated to the lower ends of the helices A and B, and when the slider is in such a position that the period of oscillation of the
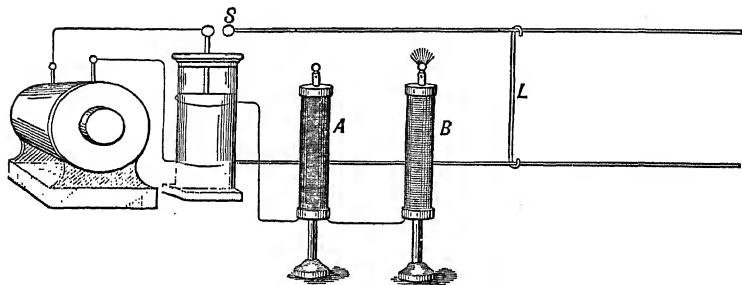
Fig. 466.　Resonance experiment

discharge is the same as the period of oscillation in the wire on A, a strong brush discharge will be observed from the upper point of that helix; while by moving the slider until the jar circuit is in resonance with B, the discharge will take place from the top of B instead of from A.

815. Electromagnetic Waves in Air. As early as 1862 Clerk Maxwell, who followed Faraday in recognizing the important function of the dielectric in all electric phenomena, showed that it was probable that when a current is stopped or started in a conductor, *the inductive action on other conductors is not communicated instantly, but is propagated through the intervening dielectric with a velocity equal to*

$$\frac{c}{\sqrt{K\mu}}$$

where μ is the magnetic permeability of the medium, K is its specific inductive capacity and c is the ratio of the C. G. S. electromagnetic to the C. G. S. electrostatic units of quantity.

The quantity $c/\sqrt{K\mu}$ can be determined by electrical experiments in a variety of ways and is found to have a value in air of very nearly 300,000,000 meters per second, which agrees with the velocity of light.

Of course, if induction is propagated with a definite velocity, an alternating current sending out first one kind of inductive disturbance and then the reverse must produce a series of electrical waves, just as a tuning-fork giving a series of impulses which travel successively forward through the air produces a train of sound waves.

816. Hertz' Experiments. Maxwell's conclusions as to electric waves were not *directly* demonstrated until 1884, when the
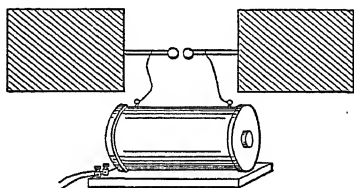
FIG. 467. Hertz oscillator   FIG. 468. Hertz resonator
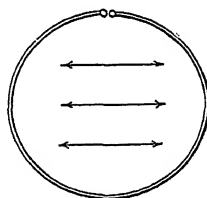
German physicist, Hertz, obtained such waves and measured their velocity.

The difficulty was twofold: to set up waves short enough to be studied — for if their velocity was 186,000 miles per second an alternating current with a frequency of even 186,000 per second would produce waves a mile long — and, *second,* to devise some method of detecting and measuring them.

Hertz succeeded in obtaining waves sufficiently short to measure by using those sent out in the oscillatory discharge of the apparatus shown in figure 467.

Two rectangular metal plates were mounted as shown, with polished knobs close together. The plates were connected to the secondary of a powerful induction coil so that when charged by the coil they discharged with oscillations across the spark gap

between the knobs. Thus a group of short waves was sent out by the oscillatory discharge every time the induction coil acted, and this may have been 200 times a second, but each group died out absolutely before the next was formed.

To detect the waves, Hertz used an *electrical resonator*, a hoop of metal having at one point a minute spark gap between two knobs. The resonator was adjusted to be in resonance with the vibrator so that in a darkened room a small spark could be seen at the spark gap of the resonator at every discharge of the vibrator, even when it was 10 or 12 meters distant.

In order to test whether the disturbance was propagated as a wave motion, Hertz set up the vibrator in front of a great

reflector of sheet metal, as shown in figure 469, so that the reflected waves meeting the advancing ones might cause nodes and loops just as in any other case of wave motion. *He then found by means of the resonator that there actually were points of maximum disturbance and points half-way between them where the effect was a mini-*

FIG. 469.   Hertz nodes and loops

mum. In this way the existence of electrical waves was proved and the wave length measured.

From the wave length and period of oscillation the velocity of the waves was calculated *and found to be, as nearly as could be determined, the same as the velocity of light,* thus confirming the anticipations of Maxwell.

817. Other Experiments. Later experimenters have devised oscillators of other forms more suitable for obtaining short waves. One of the best arrange-ments is that of Righi shown in figure 470. Two brass balls are mounted near each other, the space between being filled with oil contained in a surrounding glass cylinder. Just outside of

FIG. 470.   Righi oscillator

these are other balls connected with the poles of the induction coil. A large difference of potential between the two inner balls

is required before a spark can burst through the oil, and consequently the vibrations are so much the more energetic.

Using oscillators of this form electric waves only a few millimeters long have been obtained and measured, and have been reflected, refracted, and polarized, like waves of light.

Recently, by the use of a modification of the Righi method
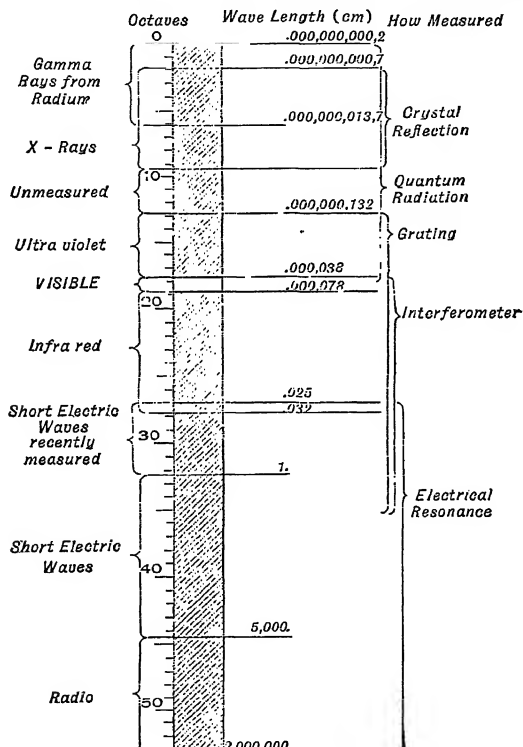
FIG. 471.   Electromagnetic spectrum

in which tiny cylinders of tungsten wire were used, Nichols and Tear succeeded in measuring the wave length of electric waves less than three mm. long, the shortest wave length obtained by Righi being about ten mm.

This work is important because it bridged the only remaining gap between electric wave and heat wave spectra.   The range of wave lengths covered by this work is shown in figure 471.

818. Radiotelegraphy. The application of electromagnetic waves to wireless or radiotelegraphy was first made by Marconi in about 1896. Figure 472 shows an old form of transmitting circuit. The waves are sent out from an *antenna* or *aerial A* which may consist of a pair of copper wires stretched at a height of from 50 to 100 ft. above the ground. The *aerial* is directly connected to the ground *E* through a coil of perhaps a dozen turns of heavy copper wire. The condenser *C* has its coatings connected by a circuit which takes a few turns close around the coil *B* in the antenna circuit, and includes a spark gap *S*. When the condenser is charged by an induction coil or high tension transformer, discharges take place across the spark gap, accompanied by oscillations or surgings in the condenser circuit, and these, by induction between the coils of wire in *B* and *D*, set up corresponding oscillations in the antenna circuit, and if one circuit is in resonance with the other the oscillations will be strong.

819. Production of Sustained or Undamped Waves. In all devices which employ spark discharges for the production of electric waves such as that just described and also in the oscillator of Hertz (§ 816) each spark discharge produces a train of electrical oscillations which rapidly die out (see Fig. 472) due to energy being expended in light, heat, etc. Oscillations which die out in this way are said to be damped. The great advances in radiotelegraphy and the development of radiotelephony in recent years are largely due to the ease with which continuous trains of *sustained* or *undamped* electric waves can be produced by the three electrode vacuum tube.

FIG. 472. Wireless sending circuit

A simple oscillating vacuum tube circuit with an aerial is shown in figure 473.

The aerial and the coil *L* make up the circuit of oscillation, the electricity surging alternately up into the aerial and down into the earth, the aerial and the earth acting as the two coatings of a

condenser separated by a wide air space. This capacity C can be calculated approximately and is called the capacity of the aerial. It is usually small compared with that of an ordinary Leyden jar.

The oscillations are built up as follows: suppose an electrical disturbance gives the grid a small positive charge as may result from simply throwing on the A or the B battery. This produces a sudden impulse of current in the plate circuit from the B battery. But the plate circuit is so connected that this impulse of current has to pass through part of the coil L, producing a corresponding inductive voltage reaction in the coil. Therefore, the potential of the grid is disturbed still more because it is connected to
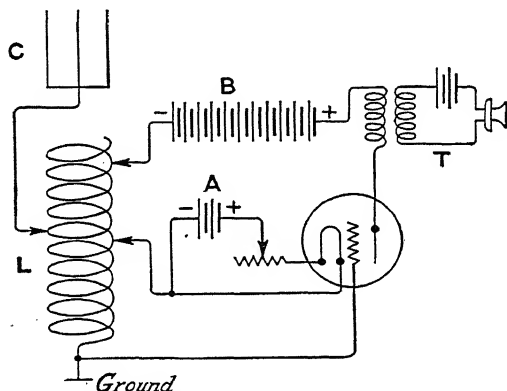
Fig. 473. Oscillating vacuum tube circuit with aerial

one end of this coil, and, furthermore, the grid is so connected that when the current in the plate circuit begins to rise the potential of the grid begins to rise at the same time. This tends to give even a further impulse to the plate current flowing through the coil. In like manner a very small fall of grid potential will result in a considerable decrease in or downward impulse to the current in the coil. But this is the condition necessary for an oscillation. When an upward impulse is given to the current in the coil a charge surges up into the aerial. It immediately recoils downward again, is given an additional downward impulse, and recoils upward again, and so on, the upward and downward surging building up because of the impulse given to each surge until the energy is *dissipated* as fast by the resistance of

the wire along which it oscillates and by the radiation of electric waves from the aerial as the energy is *supplied* by the B battery. The period of the oscillation is calculated just as in the case of the discharge of a Leyden jar (§ 811), where L is the inductance of the coil up to the contact point where the aerial is connected, and C is the capacity between the aerial and the earth. Since the period of the oscillation and the velocity of electric waves are known, the wave length is easily determined.

820. Transmission of Electric Waves Through Space. An indication of the way electric waves are transmitted from an aerial consisting of a vertical wire grounded at one end is shown in figure 474. As the aerial is charged first negatively and then positively with regard to the earth, lines of electric force are first established in one direction and then in the opposite between the aerial and the earth, a wave of electric force being thrown out sideways in all directions with each electrical oscillation. These waves are transmitted through space with the velocity of light. The lines of electric force are seen to be vertical at the earth's surface. All electromagnetic waves consist of waves of both electric and magnetic force superimposed, the lines of magnetic force always being at right angles to the lines of

Vertical Section showing Electric Field

Aerial

Earth

Horizontal Section showing part of Magnetic Field at Right Angles to Electric Field.

Fig. 474. Electric wave train

electric force. In the figure the magnetic lines of force are expanding circles parallel to the earth's surface shown in the horizontal section in the lower part of the figure. An electromagnetic wave is impossible without the electric and magnetic lines of force superimposed at right angles.

The formula for the velocity of electric waves (§ 815)

$$V = \frac{c}{\sqrt{K\mu}}$$

is exactly analogous to that of § 290 for the velocity of transmission of sound

$$V = \sqrt{\frac{E}{d}}.$$

In the former, $\frac{1}{K}$, the reciprocal of the specific inductive capacity may be thought of as the electrical elasticity of the ether, and μ, the permeability, as its electrical inertia.

The distribution of electromagnetic waves going out from an ordinary horizontal aerial is not so simple as this. Furthermore in the so-called *Kennelly Heaviside layer* a few miles above the earth's surface the presence of free electrons alters the velocity and therefore the direction of the waves, so that they are turned downward by small amounts depending on the wave length. This explains, in part, why radio waves of different wave lengths have different ranges, and are often more strongly received at greater than at small distances.

821. Receiving Circuits. A simple form of receiving circuit used in wireless telegraphy is shown in figure 475. The antenna is connected directly to earth through a cylindrical coil known as the primary coil. Another coil, the secondary, of smaller diameter, is mounted so that it can be slid inside of the primary coil or moved away from it when it is desired to weaken the inductive action of one coil upon the other. The terminals of the secondary coil are joined to the coatings of a variable condenser of small capacity.

When electric waves of a certain period fall on the antenna, oscillations are set up which will be strongest when the natural period of electrical surgings in the antenna and primary coil are the same as the period of the oncoming waves. This adjustment may be effected by regulating the number of turns used of the primary coil by means of a sliding contact, or by connecting additional turns from a so-called "loading coil." The secondary also is brought into resonance with the primary, by the adjustment of the variable condenser and distance between the two coils. Figure 476 shows another receiving circuit which operates

on the same principle as the former. The detecting system, however, is different.

822. Detectors. There are two different kinds of detectors in common use, the crystal detector and the vacuum tube detector, the former being shown in figure 475 and the latter in figure 476.

In both of the circuits shown, the detector system is connected across the terminals of the variable condenser. The receiving telephones are wound with a large number of turns of very fine wire so that they may respond to the slightest current.
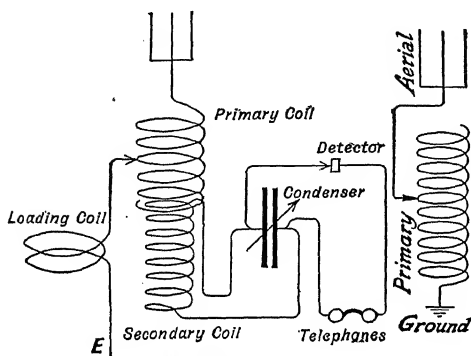
Fig. 475

A detector is necessary in radio reception because the electrical oscillations, which are received, are too rapid to affect the telephone diaphragm. The detector permits the current to flow more easily in one direction than in the opposite direction, or has the power to *rectify* so that as oscillations of varying intensity are received, a varying current in one direction passes through the telephone, which rises and falls as the high frequency oscillations increase or decrease in intensity as shown in the two lower curves of figure 477.

Fig. 476

The action of the three electrode vacuum tube as a detector is as follows: The incoming oscillations produce oscillations of charge on the grid which rise and fall with an increase or decrease in intensity of the incoming oscillations. But the grid action is such that the positive half of the corresponding plate current oscillations is greater than the negative as shown in figure 477 so that on the average a positive plate current flows which rises and falls with the rise and fall of the intensity of the high frequency oscillations. These current fluctuations are greatly amplified in energy compared with

those obtained by a crystal detector because relatively large fluctuations of energy from the B battery are produced by the slightest variations of grid potential. This fluctuating plate current is superimposed upon the steady

Incoming Oscillations

Actual Plate Current

Average Plate Current Showing Pulses
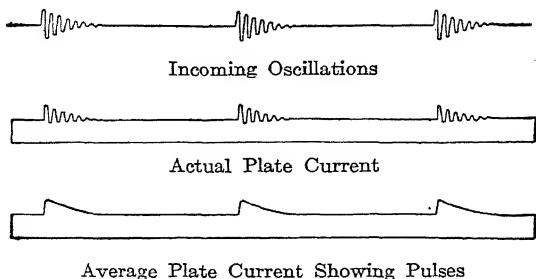
FIG. 477. Current curves in detection

current from the B battery as shown in the two lower curves of the figure. The steady current does not disturb the response of the telephone diaphragm to the fluctuating current produced by the fluctuating intensity of the incoming oscillations.

823. Amplification. The three electrode vacuum tube may also be used as an amplifier. If the grid is so connected that its potential is made to rise and fall due to electrical oscillations, even though the oscillations be very minute, comparatively strong pulsations of current from the battery B may result. The incoming wave energy, though only sufficient to cause a small pulsation of charge on the grid, acts as a trigger on the battery B, and may release large pulsations of energy from it sufficient to produce sound in the telephones which are connected in this circuit. With most radio sets the *radio frequency* oscillations are amplified before detection is accomplished as well as the sound frequency or *audio frequency* after detection. In amplifying the radio frequency, tubes are often used with a fourth electrode in the form of an extra wire grid between the plate and control grid, which is kept at a constant potential. This acts as a screen to reduce the disturbing action of the small electrical capacity between the plate and the control grid.

824. Radiotelephony. Radiotelephony requires a continuous train of waves, and it is the comparative ease with which a train of undamped waves can be produced by the three electrode tube,

which has given such an impetus to wireless telephony. By the use of a suitable transmitter these oscillations may be made to fluctuate in intensity according to the voice fluctuations just like the current in an ordinary telephone, and it is these fluctuations in intensity which act upon the receiving telephone and reproduce the sounds of the voice. In figure 473, T shows a transmitter which consists of an ordinary telephone transmitter
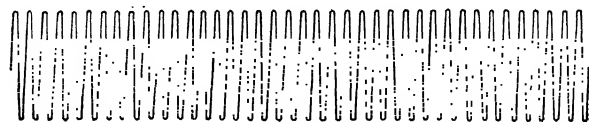
Undamped High Frequency Oscillations

Oscillations Modulated According to the Vibrations of the
Diaphragm of the Telephone Transmitter

Average Current Fluctuations in Telephones of Receiver
Showing the Correspondence with the Modulations
of the Incoming Wave Train

FIG. 478.   Current curves in wireless telephony

whose fluctuating current, flowing through the primary coil of a small transformer, produces corresponding fluctuations in the potential of the plate due to the action of the transformer secondary coil, to one end of which the plate is connected. In the vacuum tube variations in the plate voltage produce proportional variations in the intensity of the oscillations in the aerial. The intensity of the high frequency wave train sent out from the aerial then fluctuates in close correspondence to the sound frequency fluctuations of the voice received in the telephone transmitter.

For instance, suppose the transmitting station sends out a continuous series of undamped waves 1000 meters long. These will come along at the rate of 300,000 to the second, and if the receiving station is in resonance with these waves, a strong electrical oscillation is set up, represented by the upper curve of figure 478, but these oscillations are far too rapid to set up vibrations in the telephone diaphragm. Now if sound waves acting through the transmitter cause fluctuations or variations in intensity in the original waves, the effect will be such as indicated in the middle curve. This also would not directly affect the telephone, for the positive and negative currents exactly balance each other. A detecting device has to be used which may be in the form of a crystal detector, or if a more sensitive detector is required, an audion detector may be used. Then the current in the telephones will on the whole vary as shown in the lower curve, corresponding to the fluctuations of the incoming wave train, and the diaphragm will vibrate in response, thus reproducing the sound.

This varying of the intensity of the continuous train of oscillations which are sent out is called *modulation*. It may be accomplished in different ways, usually through varying the plate voltage. The method described is a very simple way of modulating through plate voltage variation.

LIGHT

SHADOWS AND PHOTOMETRY

825. Light. In a perfectly dark room we cannot see any objects — we have no sensation of sight — showing that vision requires something more than simply the eye and the object to be seen. That additional something is called *light*. We may define it as the agent which excites the sensation of sight.

When a candle is lighted in the room we see it and also the other objects near. The candle flame is said to be self-luminous and a source of light.

Conditions of Vision. When the candle is so screened that its light falls only on the eye of the observer but not on other objects in the room, then only the candle itself is seen. It thus appears that illuminating the eye does not give it power to see other objects from which light is excluded. *Light must fall upon the objects themselves if we are to see them.*

And even when an object is illuminated, if a screen is interposed across the straight line from the object to the eye, the former is hidden and we see the screen but not the object behind it. This leads to the inference that *in order that a body may be seen light must pass from it to the eye, and usually this takes place along straight lines.*

826. Transparent and Opaque Bodies. Bodies differ greatly in their capacity for transmitting light. Those that transmit it freely are said to be *transparent*, while those that intercept it are called *opaque*. Opaque bodies are of two kinds: those that turn back the light at the surface and those into which light penetrates and is absorbed and transformed into heat. The opacity of metals is largely of the first kind, while that of most other substances is due to absorption.

Substances like paper or milk-glass, or milky or muddy waters, which transmit light but through which we cannot see objects, are said to be *translucent*. They are not homogeneous bodies, but light in passing through them is scattered in all directions at the surfaces of innumerable little particles throughout the

mass. Even transparent bodies, such as glass or water, turn back or reflect *at the surface* a part of the light that falls upon them.

Bodies may be transparent for some kinds of light and opaque for others, and this is largely the cause of the colors of bodies.

827. Light Advances in Straight Lines. *Light travels out from the source in straight lines so long as it remains in a homogeneous medium.*

Thus a carpenter sights along the edge of a board to see whether it is straight; and the boundary of a shadow is roughly
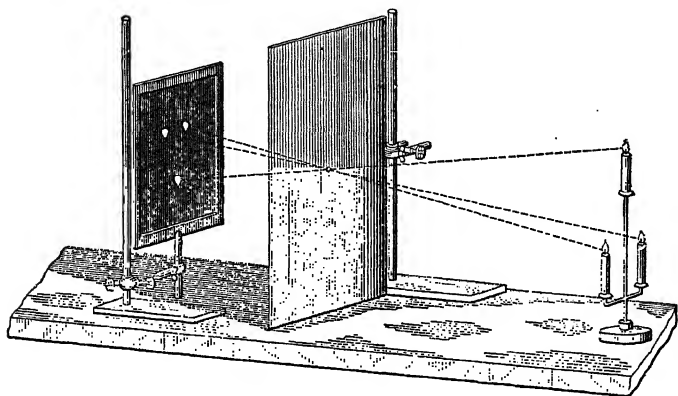
FIG. 479. Light travels in straight lines

defined by straight lines through the source of light and tangent to the obstacle.*

But *the most convincing evidence of this fact is the exactness with which surveys are made.* All measurements of angles made by surveying or astronomical instruments assume that light from the distant object comes to the observer's telescope in straight lines if the medium is homogeneous.

When light is admitted to a dark chamber through a small hole, an inverted image of the outer landscape is formed on a white screen opposite the opening. For as light goes through the opening in straight lines each point on the screen is illumi-

* On close examination of a shadow, even when the source is the merest point of light, it is seen that there is not a sharp transition from light to dark at its edge, but it is marked by a series of alternate dark and light *diffraction bands*.

nated by light from a single point in the landscape, and, therefore, the relative brightness and colors of objects in the landscape are reproduced at the corresponding points on the screen. A white screen is used because it reflects to the eye all colors equally well.

828. Shadow, Umbra and Penumbra. When the source of light is a broad luminous surface, as in case of an ordinary gas

FIG. 480.   Umbra and penumbra

flame, shadows are not sharply defined, but shaded at the edges. For example, in case of the sun and earth, as shown in figure 480, the region between B and C is in full shadow, and is known as the *umbra;* while the outer region shades from full illumination at A to complete shadow at B, and is known as the *penumbra.*

829. Intensity. If the source of light S, figure 481, *is a point*, it is clear that a surface A if moved to B, twice as far from the source, will intercept only one-fourth as much light as in its original position; if its distance from the source is increased three times it will intercept only one-ninth as much light, etc. Hence *the intensity of its illumination or the quantity of light*

FIG. 481

which it receives per unit surface varies inversely as the square of its distance from the source.
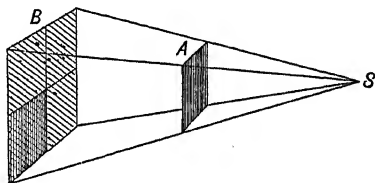
That is, if the intensity at unit distance from a point source is I_0 then the intensity I at a distance r from the source is

$$I = \frac{I_0}{r^2}.$$

830. Oblique Incidence. Suppose a square surface is placed perpendicular to the rays of light, as shown at AB in figure 482, and is illuminated with light of intensity I. If it is inclined through an angle x into the position AC, the beam of light falling upon it will be narrower than before in the ratio AD to AB. The intensity of illumination in the inclined position will there-

fore be to the intensity when perpendicular in the ratio of AD to AB, that is, as *cos x* is to 1.

831. Actual Sources. In all practical cases the source is not a point, but a luminous surface or region from every point of which light is sent out; and it is clear that however close the illuminated surface may be placed, its intensity of illumination cannot be greater than the brightness of the source.

Fig. 482

The law that the intensities at two points are inversely proportional to the squares of their distances from the source holds very closely when they are in the same direction from the source and are so far away from it that it subtends only a small angle. If the source subtends an angle of 10° the error is less than 1 per cent.
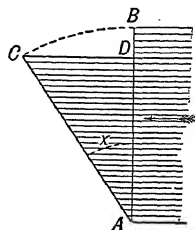
832. Photometry. The measurement of the relative amounts of light given out by two sources is called *photometry*. When both lights are of the same color the comparison may be made by several methods with much accuracy. If the lights differ in color they may be analyzed into their component colors and the corresponding components of each be compared by means of a spectrophotometer to be described later (§ 947).

In all photometric measurements care must be taken that the only light falling on the photometer screen comes *directly* from the lights which are compared. The experiments must therefore be conducted in a room from which daylight is excluded, and by means of black screens all light reflected from white walls or other objects must be kept from the photometer screen.

833. Rumford Photometer. The simplest form of photometer is that devised by Count Rumford and shown in figure 483. An opaque rod is mounted a short distance in front of a white screen. The lights to be compared are so placed that the two shadows of the rod are side by side and of equal intensity. The shadow cast by A is illuminated only by B and that cast by B is illuminated only by A; if therefore the shadows are equally intense the illumination of the screen must be the same by A as by B. Let b represent the brightness of A, meaning by brightness the intensity with which it illuminates a surface at

unit distance from it, and let b_1 represent that of B and let d and d_1 represent their respective distances from the screen. Then

$$\frac{b}{d^2} = \frac{b_1}{d_1^2} \quad \text{or} \quad \frac{b}{b_1} = \frac{d^2}{d_1^2}$$

or the brightnesses are proportional to the square of the distances measured from the lights to the screen.
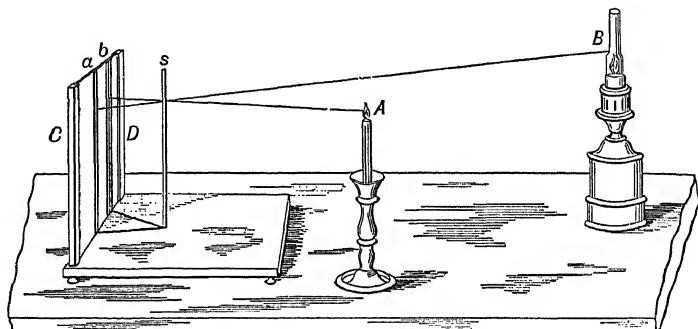
FIG. 483.   Shadow photometer

334. Bunsen Photometer. A better form of photometer, free from penumbral disturbances at the edges of shadows, is the grease-spot photometer devised by Bunsen, which consists of a screen of white paper having a spot at its center rendered
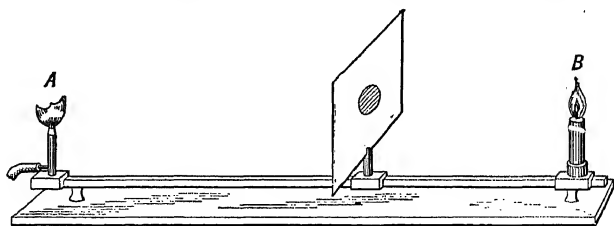
FIG. 484.   Grease-spot photometer

translucent by means of grease or paraffin. The screen is placed between the lights to be tested so that one side is illuminated by one light and one by the other. The translucent spot transmits light quite freely, and therefore if the paper is lighted only on one side the illuminated side will appear bright with a dark spot at the center while the side away from the light

will be darker with a bright central spot. If the two sides of the paper are equally illuminated the spot disappears. The intensities of the lights are then proportional to the squares of their distances from the screen.

835. Lummer-Brodhun Photometer. The Lummer-Brodhun photometer is a development of the idea embodied in the Bunsen photometer. Its construction is shown in figure 485. At S is an opaque screen with perfectly white plaster-of-Paris surfaces. M and M' are mirrors and P and P' are two polished prisms of glass, one of which, P', has the whole of the diagonal surface completely polished, the other, P, has only a central round spot on the diagonal face polished, the rest being ground away so that the two prisms touch only in this central polished spot. When freshly polished they are put together under great pressure so that they cohere firmly like a solid block of glass. If the lights to be compared are now placed in the direction of the arrows A and

FIG. 485

B, respectively, light from the side of the screen illuminated by A after reflection at M passes directly through the central spot in the block of prisms to the observer at O, while light from the side toward B after reflection at M' is reflected to the observer at O from that part of the diagonal face P' which surrounds the central spot, while the light from M' which falls upon the central spot simply passes through and is not reflected to the eye. Thus to the observer at O the brightness of the central spot depends on the illumination of the left-hand side of the screen by A, while the brightness of the surface around the central spot depends on the illumination of the right-hand side of the screen by the light B. The distances of the lights from the screen are varied until the central spot and surrounding surface appear equally illuminated.
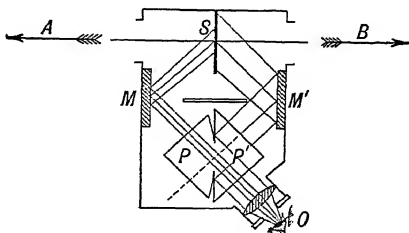
836. The Rood-Whitman Flicker Photometer. When two lights differ in color it is difficult to compare their intensities by the above methods, for the two parts of the photometer screen cannot be made to look alike. In such cases the relative luminous intensities may be approximately found by the flicker photometer. In this instrument light from the source B falls on a white screen F, fixed at an angle of 45° in front of the eye tube through which it is observed. Light from the source A falls on a second white screen which also makes an angle of 45° with the eye tube and is rotated slowly about the axis shown in the figure.

This screen is made with projecting sectors which, as it rotates, come between the eye tube and the fixed screen F, so that the observer sees during one-quarter of a revolution only the rotating sector illuminated by the light

A, while in the next quarter revolution the fixed screen illuminated by *B* is exposed.

Thus the two screens are alternately exposed to view for equal times, and by careful adjustment the speed of rotation is made such that a very disagreeable flickering effect is noticed unless the illuminations due to *A* and *B* are of equal intensity. The distances of *A* or *B* are varied until the flickering is a minimum, when the intensities of the two lights are proportional to the squares of their distances from the screens.

Plan of Rotating Screen

Eye
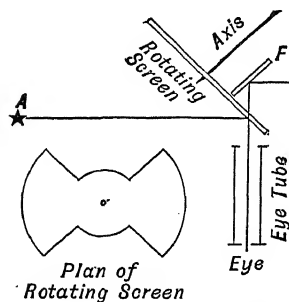
FIG. 486

837. Standard of Light Intensity. Many standards of light have been proposed for commercial and scientific purposes, but none are altogether satisfactory.

The English Electrical Standards Committee has defined one candle-power as one-tenth the candle-power under standard conditions of a particular pentane lamp kept in the National Physics Laboratory. The standards of light used by our Bureau of Standards are certain incandescent lamps which have been compared directly or indirectly with the English, French and German standards.

The English standard candle formerly used was a spermaceti candle made to burn 120 grains per hour with a flame height of 45 mm.

838. Illumination. The illumination of a surface is measured in foot-candles, one foot-candle being the illumination produced by a 1-candle-power lamp at a distance of 1 ft., or by a 16-candle-power lamp at a distance of 4 ft.

Some Values of Illumination

Good illumination for reading................ 4 foot-candles.
Poor illumination for reading................ 1–2 foot-candles.
Full moonlight............................. .02 foot-candles.

VELOCITY OF LIGHT

839. Early Experiment. It has been seen that light seems to pass from the source to the eye in straight lines. This at once suggests inquiry whether or not the eye *instantly* experiences the sensation of light when a candle is uncovered.

The earliest attempt to solve this question was made by the Florentine Academy after a method proposed by Galileo. A light on an eminence was uncovered and flashed to a station on a distant hill where a second observer also having a covered light was watching. As soon as the flash was seen by the second observer he uncovered his light, sending an answering flash back to the first station. The first observer was to note the exact time between the uncovering of his light and the sight of the return flash. The experiment showed that if any time at all was required for light to travel from one station to the other it was too short to be detected by that method.

840. Roemer's Discovery. *The first evidence that light required an appreciable time to pass from one point to another*
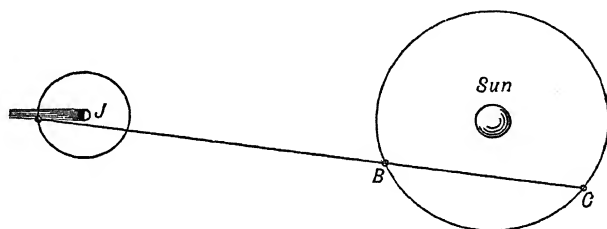
FIG. 487

was obtained by the Danish astronomer, Roemer, in 1676, *by the* following method:

The first satellite of Jupiter passes into the planet's shadow and disappears or is eclipsed every time it revolves around the planet. Some years before Roemer's discovery Cassini had carefully determined the periodic time of the satellite and had prepared tables showing when the eclipses might be expected to take place for several years ahead. On comparing these tables with the recorded times of observed eclipses Roemer found that they were observed sooner than predicted when the earth was on the side of its orbit nearest to Jupiter, and later than predicted when it was on the opposite side. He concluded that the discrepancy was due to the velocity of light; for evidently if it takes 10 minutes for light to cross the earth's orbit from B to C, then an eclipse would be seen 10 minutes later if the earth were at C than if it were at B. The observations indicated

that light requires 16 minutes to cross the whole of the earth's orbit, or approximately 8 minutes to go from the sun to the earth or, more exactly, 498 seconds to traverse the 92,900,000 miles between sun and earth, making the velocity of light in inter-planetary space 186,600 miles or 300,200 kilometers per second.

841. Bradley's Discovery.   No further evidence of the velocity of light was obtained until 1727 when the English astronomer, Bradley, discovered that *the stars in any given part of the heavens were apparently displaced from their mean positions by an exceedingly small amount which depended on the position of the earth in its orbit*.   The explanation of this phenomenon, which is known as *aberration*, was finally suggested to him by the observation that the position of a flag on a small boat depended on the velocity and direction of motion of the boat as well as on the wind.   He said to himself that the apparent direction in which light comes to the earth from a star must be affected by the velocity of the earth, just as the apparent direction of a breeze to a man in a boat depends on the motion of the boat.
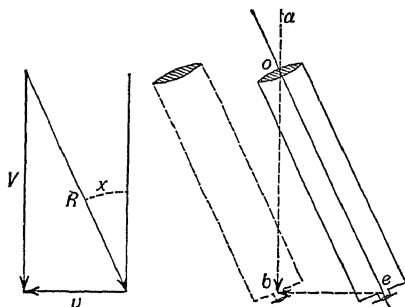
FIG. 488

For suppose light coming from a star in the direction *ab* (Fig. 488) enters at *o* a telescope which is being carried along sidewise in the direction *eb*, and that the light advancing in the direction *ob* reaches *b* at the same instant that the eye-piece reaches there as it moves from *e* to *b*.   The light will be received by the eye and the telescope will seem to be pointing at the star.   To accomplish this the telescope evidently must incline forward so that $ob : eb :: V : v$ where V is the velocity of light and v is the component, perpendicular to the star's direction, of the velocity with which the telescope is carried along by the earth, and the apparent direction of the star differs from its true direction by the angle x, such that

$$tangent\ x = \frac{v}{V}.$$

When the earth is moving directly toward or away from a star there is no displacement or aberration, while stars in directions at right angles to that in which the earth is moving have maximum displacement. The apparent position of a star therefore changes slightly as the earth moves from one part of its orbit to another, so that by careful determinations of its apparent position made during an entire year the maximum displacement or *aberration constant* may be determined.

Recent observations give as the aberration constant 20.492″. Now, the mean velocity of the earth in its orbit is 18.51 miles per second, and we may calculate the velocity of light V from the relation

$$tan\ (20.492'') = \frac{18.51}{V}$$

which gives 186,400 miles or 299,930 kilometers per second.

842. Fizeau's Method. On account of the enormous velocity of light it was not until 1849 that a method of measuring it was devised which did not involve astronomical measurements. In that year the determination was made by Fizeau by the following method. A telescope and collimator were set up 8.633 kilometers (more than 5 miles) apart. A beam of sunlight L (Fig. 489) sent through an opening in the side of the telescope was reflected by a small oblique plate of glass G so that it passed directly out through the lens of the telescope to the distant collimator which was provided with a mirror M at its back. The collimator and mirror were so adjusted that the beam of light was reflected directly back into the telescope again, and passing *through* the plate of glass G was received by the eye at E. Thus light came to the eye at E after traveling to the distant mirror M and back again. At S in the telescope is a small opening which is alternately opened and closed by the teeth of a cogged wheel W which revolves immediately in front of it. If the wheel is slowly rotated, light from L passing through a gap between two teeth travels to the distant mirror and back again through the same opening to the eye at E. If the notches and teeth in the wheel are of equal width and if the speed of rotation is such that a tooth moves forward just its own width in the time that light requires to go to the distant station and back, light which

must have passed out through an opening will on returning find the opening closed by a tooth and will therefore be cut off from the observer at E. If the speed is then doubled, light passing out through one opening will return through the next one; at a still higher speed it will be eclipsed again, etc. It is therefore only necessary to observe the speeds at which the light is completely eclipsed to be able to determine the velocity of light, when the distance between the two stations and the number of teeth in the wheel are known. In Fizeau's apparatus there were
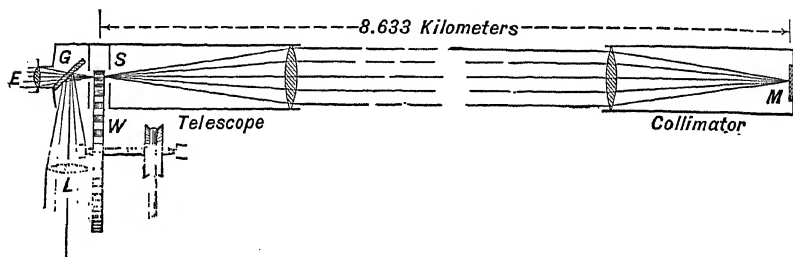
FIG. 489. Fizeau's apparatus for measuring the velocity of light

720 teeth in the wheel and the first eclipse was noticed when the wheel made 12.6 revolutions per second. Therefore the time required for light to travel twice the distance between the two stations was only

$$\frac{1}{720 \times 12.6 \times 2} - \frac{1}{18143} \text{ sec.}$$

The stations were 8.633 kilometers apart, making the velocity of light 313,000 kilometers per second. The same method carried out by Cornu in 1874 with improved apparatus gave $V = 304,000$ kilometers per second.

843. Foucault's Method. Another method of measuring the velocity of light was devised and carried out by the French physicist, Foucault, in 1850. The essential features of the apparatus are shown in the diagram, figure 490. A beam of sunlight concentrated on the narrow slit S passes through it and through the inclined plate of glass G and the lens L to a small mirror m, from which it is reflected to a concave mirror M whose center of

curvature is exactly at the center of the mirror m. The light is reflected perpendicularly back from M to m and thence back to the glass plate G, which reflects it aside into the eye-piece E. A bright image of the slit S is formed at a by means of the lens L, and this is seen by the observer at E. If the mirror m is now slowly rotated in the direction of the arrow, the image of the slit will be formed at a only when m is in such a position, as it revolves, that the beam of light reflected from it meets the concave mirror M; consequently the image at a will disappear and reappear once in each revolution; but as the speed is in-
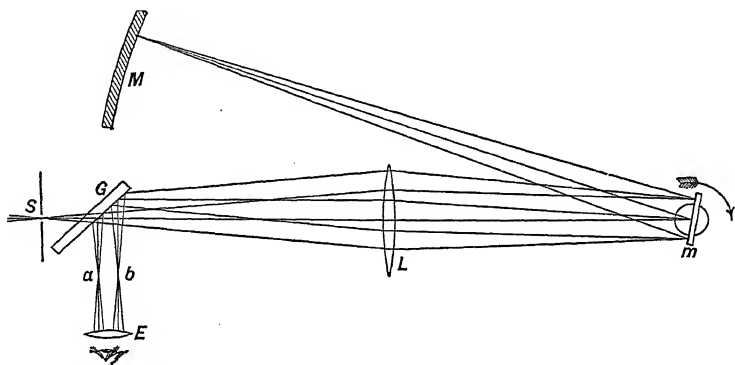
FIG. 490.   Velocity of light measured by rotating mirror

creased to more than about 10 revolutions per second the eye no longer detects the intermittence, but sees a continuous image of the slit. As the speed of rotation increases, it is noticed that the image of the slit is no longer at a, but is displaced toward b, the amount of the displacement being proportional to the speed of rotation of the mirror. This displacement is due to the fact that while light is traveling from m to M and back again, the mirror has turned forward through a small angle, and consequently the returning light is reflected slightly upward, as shown in the above figure, and not back along its original path, causing the image of the slit to be displaced to b. The displacement ab may be measured by a micrometer, and from it the angle through which the returning beam is turned upward may be determined. But this angle will be twice the angle through which the mirror turns while light is traveling from m to M and back (§ 855). All

that remains therefore is to determine the distance mM and the speed of revolution of the mirror; the velocity of light may then be easily calculated.

In Foucault's experiment the distance mM was only 4.12 meters, and the greatest displacement ab was about 0.3 mm. when the rotating mirror was making 800 turns per second, a displacement too small to give a very accurate result. But he tried the very important experiment of introducing a long tube of water between m and M through which the light was sent, and was able to show that *the velocity of light in water was less than in air*, a result of the greatest significance in determining the nature of light (§ 877).

844. Michelson's Modification. In 1879 Michelson, then at the United States Naval Academy, modified Foucault's method by substituting for the concave mirror M a lens through which the light passed to a distant mirror where it was reflected back. In his first experiments the distance from the revolving mirror to the fixed mirror was 605 meters. This great increase in the distance between the mirrors caused a correspondingly greater displacement which could be measured with far less percentage of error. His experiments in 1879–82 and those conducted according to his method by Newcomb in 1882 are the most accurate determinations of this important constant that have been made.

In some of Michelson's experiments the displacement to be measured by the micrometer was 13.3 cms., or 400 times that obtained by Foucault.

The results obtained by this improved method are as follows:

OBSERVER	KILOMETERS PER SEC.	MILES PER SECOND
Michelson, 1879...............	299,910	186,380
Michelson, 1882...............	299,853	186,345
Newcomb, 1882...............	299,810	186,317

845. Velocity Same for All Colors. In these various determinations of the velocity of light, the light employed was either sunlight, starlight, or light from the electric arc. But although these lights are complex, there was not found in any case a

perceptible difference in velocity between light of different colors, when measured in air or in interplanetary space. Shapley has shown that the velocity of light of different colors as measured from the spectra of variable stars agrees to one part in a billion.

PROBLEMS

1. How far from a screen must an 8 candle-power lamp be placed to give the same illumination as a 16 candle-power lamp 10 ft. distant?

2. When a photometer screen is equally illuminated by a 32 candle-power lamp at a distance of 2 meters, and an arc lamp 12 meters away, what is the candle-power of the arc light?

3. A rifle bullet has a speed of 2000 ft. per sec. How many inches will it advance while light travels a mile, and how far while light travels 25,000 miles (the circumference of the earth)?

WAVE THEORY

846. Mode of Propagation. Only three methods are known by which energy may be transmitted from one point of space to another. *First, by the movement as a whole of some medium reaching from one point to the other*, as in the case of ropes, belts, or shafting.

Second, by projectiles, as in the case of a shot from a gun or a ball thrown.

Third, by waves, as in case of sound or water waves.

We have found that light is communicated from one point to another with a velocity of about 300,000,000 meters or 186,400 miles per second. By which of the above processes is it propagated? We may evidently reject the first as inconceivable. The second was advocated by so great a philosopher as Sir Isaac Newton, while Huygens, the celebrated Dutch physicist, urged the claims of the third. While much of the most convincing evidence will be found in phenomena that must be taken up later in our study, there are some considerations which even at this point may help us to reach a tentative conclusion.

The velocity with which a projectile travels depends on the initial impulse. If light is communicated by means of particles shot out from the luminous body we should expect to find the velocity depending on the source and that particles emanating

from the sun would have a different velocity from those from an electric light.

On the other hand, *the velocity of a wave depends only on its wave length and the nature of the wave (whether compressional or transverse, etc.) and the properties of the medium of which it is a disturbance.* Sound waves from fiddle, pipe, or drum advance with the same speed through air. If light is a wave motion we may expect to find light waves, whatever their source, traveling with the same velocity through space, and this is precisely what experiment shows to be the case. This consideration therefore points to its being a wave motion.

847. The Ether. On the other hand, if light is propagated by waves, they are waves of what? Light passes through interstellar space and through the most perfect artificial vacua that can be produced. If there are light waves, they must be in some medium which extends throughout space as far as the most distant star from which we receive light, it must fill all vacua and permeate all bodies through which light can pass. And yet no resistance to the motion of earth or planets through this medium has ever been detected. Yet in spite of these objections such a medium must be supposed to exist if light is communicated by waves, and it has been named *the luminiferous ether* or simply *the ether.*

848. Other Evidence for the Ether. It is remarkable that there is independent evidence for the existence of such a medium obtained from the study of electricity and magnetism. Electric and magnetic forces act through vacua, and may be produced as Faraday supposed by tensions and pressures in a surrounding medium. When a magnet draws a piece of iron to itself we may in imagination see it pushed up to the magnet by the stresses in the ether.

But far more important is the direct evidence of Hertz's experiments; for *electric waves have been proved to exist and are found to have the same velocity as light.*

849. Electromagnetic Theory. And since the velocity of a wave depends both on the properties of the medium and on the kind of wave motion, it is highly probable that the vibrations in light waves are exactly the same as in electric waves; or, in other words, that *light waves are electric waves.*

This theory of the nature of light waves is known as the *electromagnetic theory of light;* it was proposed and developed by Maxwell in 1865. Its conception and establishment, next to that of the conservation of energy, is the most remarkable achievement of physical science in the nineteenth century.

In our further study we shall endeavor to test the probability of the hypothesis that light is a wave motion by inquiring whether it affords simple and natural explanations of the various phenomena as they arise.

850. Form of Light Waves. If light comes from a source as a series of waves, the form of a wave, as it advances in all directions with equal velocity, must be spherical, and the direction of advance being radial is at right angles to the wave front.

851. Beams and Rays. When light shines through a small opening, the stream of light is called a *beam;* and a very narrow beam is called a *ray.* When the beam comes from a very distant source, the rays of which it may be conceived as made up are parallel, and it is called a parallel beam; in that case the wave fronts are planes.
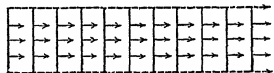
FIG. 491. Parallel beam with plane waves

When light comes from a point, the rays diverge radially from the source and the wave fronts are spherical segments having the source as their center. Such a beam is divergent, and its waves enlarge as they advance.

By means of a lens or curved mirror, a beam of light may be made to converge toward a point which is called the focus, in
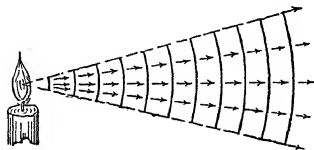
FIG. 492. Divergent beams with convex expanding waves
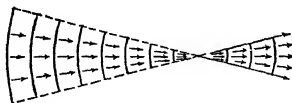
FIG. 493. Convergent beam with concave contracting waves

which case the wave fronts must be concave spherical surfaces which contract as they approach the focus.

852. Geometrical and Physical Optics. The study of light is also called *optics.* The method of treating the subject which

ignores the existence of waves and treats a beam of light as a bundle of rays is called *geometrical* optics, while the other method which investigates the dependence of the various phenomena of light on the properties of waves is known as *physical* optics.

REFLECTION OF LIGHT AND MIRRORS

853. Reflection: Regular and Diffuse. When light reaches a surface where there is a change of medium, some is reflected or turned back into the first medium while some penetrates into the second medium.

When the reflection takes place at a flat polished surface, light comes to the eye as though directly from the distant objects

themselves, and if the polish is perfect none of the light seems to come from the reflecting surface, but we seem to be looking through an opening at objects beyond. This is known as *regular reflection.*

If the surface is now ground with coarse emery, we no longer see reflected objects, but light goes out from the surface itself in all direc-
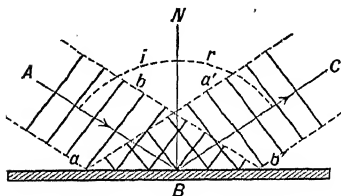
FIG. 494.   Regular reflection of waves

tions as though it were a source of light. This is known as *diffuse reflection.* It takes place at the surface of such bodies as wood, paper, cloth, etc., and seems to be due to the breaking up and scattering of light waves by the roughness or irregularity of the reflecting surface. To polish a surface so that it reflects like a mirror the very finest emery and polishing rouge must be used, a fact which indicates that the length of light waves must be extremely small. In some conjurers' illusions advantage is taken of the invisibility of a well-polished mirror surface.

854. Regular Reflection. The law of regular reflection is the same for light waves as for other forms of wave motion. If *AB* is the incident ray and *BC* the reflected one, the angles *i* and *r* which they make with the normal *BN* are called the *angles* of *incidence* and *reflection*, respectively. In case of regular reflection, *the angles of incidence and reflection are equal and lie in the same plane.* This plane is called the *plane of incidence.*

It will be observed that by reflection a wave front such as *ab* is turned into the position *a'b'*.

855. Mirror Turned Through an Angle. If a beam of light *SO* (Fig. 495) meets the mirror perpendicularly, it is reflected directly back on its path. But if the mirror is turned through the angle *x*, the reflected beam will take the direction *OP* where the angle of reflection *y* is equal to the angle of incidence *x*. By the motion of the mirror through the angle *x*, the reflected beam has therefore been turned through an angle $x + y = 2x$.
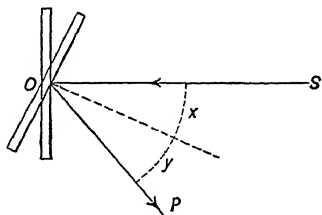
FIG. 495

Suppose the mirror is attached to the needle of a galvanometer and reflects the light from an incandescent lamp upon a graduated scale. It is clear from the above that when the needle turns through a small angle, the reflected beam of light must move through twice that angle.

856. Plane Mirror and Images. If a source of light is placed at *O* in front of a plane mirror *MM'*, figure 496, the light after reflection will appear to come from a point *O'* as far behind the mirror as *O* is in front of it. (See also § 301.) For trace the incident and reflected rays *OB* and *BC*, and produce the latter backward meeting the perpendicular *OP* at *O'*. Then the triangles *OPB* and *O'PB* have right angles at *P*, and the side *PB* common, and the angles *PBO* and *PBO'* are
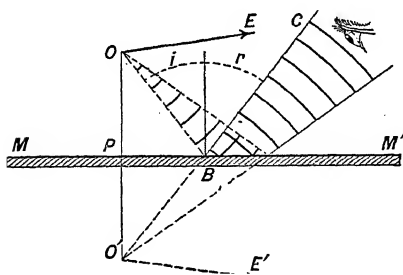
FIG. 496.   Image by plane mirror

equal because of the law of reflection, therefore $PO' = PO$. But *B* is *any* point in the reflecting surface, therefore all reflected rays from *O* if produced backward will pass through the point *O'*. Light waves from *O* after reflection from the mirror come to the eye as if they had come from *O'*, and consequently *O' is said to be the image of O. It is a virtual image as distinguished from a real one, because the light does not actually pass through the point O'.*

If an object represented by the arrow OE is placed in front of a plane mirror the image is virtual and in the position $O'E'$, each point in the image being as far behind the mirror as the corresponding point in the object is in front of it.

857. Multiple Reflection. When two plane mirrors are placed at right angles to each other as in figure 497, three images are formed of an object O placed between them. O' is the image of O in AB, O'' is the image of O in BC, while O''' is the image of O' in BC or of O'' in AB. The rays which come to the eye as if from O''' are reflected twice, once by each mirror.

In the kaleidoscope three narrow strips of mirror glass are placed edge to edge, forming a triangular prism with the mirror faces turned inward. An observer looking in at one end of the contrivance sees a regular hexagonal pattern formed by the repetition of some figure formed by pieces of colored glass at the other end of the tube.

If two flat mirrors are placed parallel and facing each other, an observer standing between them with a lighted candle will see an infinite series of images of the candle stretching into the distance in each mirror. The first image is formed by a single reflection, the second by light reflected twice, once in each mirror, etc.
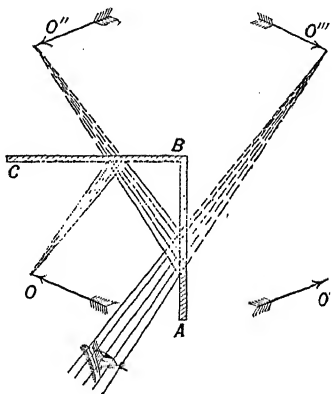
Fig. 497

858. Concave Mirror. The surface of a concave mirror is commonly a portion of a sphere, because spherical surfaces are ground and polished with comparative ease.

If such a mirror is held facing the sun, a bright spot of light, which is the image of the sun, is formed half-way between the mirror and its center of curvature. The angular size of the focal image as seen from the mirror is the same as that of the sun itself, so that the shorter the radius of curvature of the mirror, the smaller this image is. The point where the image is formed is called the *principal focus* of the mirror (Lat. *focus*, a hearth). With a large mirror of short focal length so great a concentration of the sun's rays may be obtained that lead may be melted and paper and wood ignited at the focus.

If a candle is held at A, between the center of curvature of the mirror and its principal focus, a real image of the flame will

be formed at a certain point B beyond C. This image, being real, may be seen on a white screen placed there. It is inverted and as much larger than the object at A as it is farther from the mirror.

When the candle is moved away from the mirror, the image moves toward it and they meet at C where the image is of the same size as the object and still real and inverted. As the candle is moved still farther from the mirror, the image approaches F *as a limit* and when the distance of the candle is many times the radius of curvature
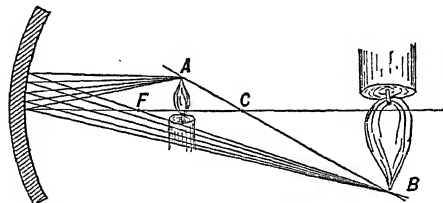
FIG. 498.   Image by concave mirror

of the mirror, the image is formed almost exactly at F, the size of the image being smaller the farther off the candle is placed.

859. Conjugate Foci. The positions of candle and image, A and B, are interchangeable — the candle may be placed either at A or B and the image will be formed at the other point. *Two points so related that one is the image of the other are known as conjugate foci.* The *principal focus* is conjugate to a point on the axis infinitely distant from the mirror.

860. Principal Focus. The principal focus of a mirror may be defined as *that point where all rays parallel to the axis meet after reflection, it is half-way between the mirror and its center of curvature.* This may be proved as follows:
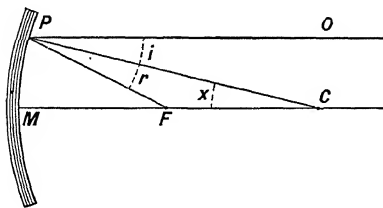
FIG. 499

Let C be the center of curvature of the mirror, and MC its axis. A ray OP parallel to the axis MC and meeting the mirror at P will be reflected into the direction PF, such that the angle of reflection r is equal to the angle of incidence i. But since OP is parallel to the axis, the angles i and x are equal, and consequently $r = x$, and the triangle FPC is isosceles and $PF = FC$. If the point P is not too far from M, PF and MF are very nearly equal, so that $MF = FC$, and F is therefore half-way between M and C.

It is clear from the above that *all* rays parallel to the axis of a concave spherical mirror do not meet exactly at the same point after reflection. This imperfection is known as aberration. When a concave mirror is only a very small portion of a sphere this aberration is slight.

861. Construction of Image. The size and position of the image which a concave mirror forms of an object in front of it may be determined by the following construction:

Suppose it is required to find the image of the arrow *PO* (Fig. 500). *Trace two rays from P, and the point where they intersect after reflection is the image of P.* One ray easily traced is *PA* through the center of curvature *C*. This ray meets the mirror perpendicularly and is reflected back along the same line *ACP*. Another ray to be taken is *PB*, which is parallel to the axis and is, therefore, reflected back through the principal focus *F*, and intersects the first ray at *P'*. This point is therefore the image of *P*. The line *PCP'* through the center of curvature is known as the *secondary axis* through *P*.
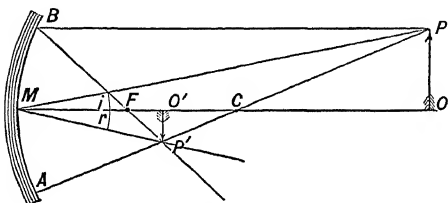
FIG. 500

The image of the point *O* on the axis will be at *O'*, also on the axis, so that *P'O'* will be the image of the arrow *PO*. This image is evidently *inverted*, it is also *real*, for rays of light from various points in the object *PO actually pass through* the corresponding points in the image *P'O'*.

862. Size of Image. *The size of the image is to the size of the object as their distances from the mirror.* For if we draw the rays *PM* and *MP'* reflected at *M*, the angles *i* and *r* are equal by the law of reflection, hence the triangles *POM* and *P'O'M* are similar and *PO : P'O' :: OM : O'M*.

It is also evident from the construction that the sizes of object and image are proportional to their distances from the center of curvature *C*.

863. Virtual Image. When the object is moved nearer to the mirror than the principal focus *F*, an *erect virtual* image is formed back of the mirror, as will be clear from the following construction. Trace as before two rays from *P*, one parallel to the

axis and reflected through F, the other perpendicular to the mirror and reflected through C; they will *diverge* after reflection and must be produced backward to find the point of intersection P'. This is the image of P, and is *virtual* because the light from P does not actually pass through P'. The sizes of object and image are
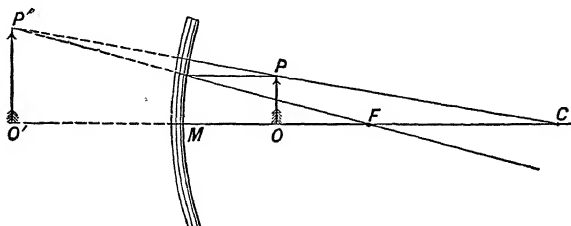
FIG. 501. Virtual image by concave mirror

proportional to their distances from C, hence the virtual image is larger than the object. As the object is moved toward the mirror the image also approaches it and they meet at M.

864. Formula for Concave Mirror. A simple formula which expresses the relation between the radius of curvature of a mirror and the distances from it of two conjugate foci, may be obtained as follows: By similar triangles (Fig. 502).
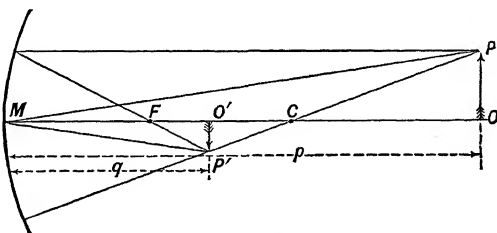
FIG. 502

$$OP : O'P' : : OC : O'C$$

also

$$OP : O'P' : : OM : O'M$$

therefore

$$OC : O'C : : OM : O'M. \tag{1}$$

Let p and q be the distances from the mirror of O and O' respectively, and let r be the radius of curvature of the mirror,

then $OC = p - r$, $O'C = r - q$, $OM = p$, $O'M = q$, and we have by substituting in (1)

$$p - r : r - q :: p : q$$

and multiplying means and extremes

$$pr - pq = pq - qr.$$

Dividing through by pqr, we obtain *the mirror formula*,

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{r}. \qquad (2)$$

When O is at a great distance from the mirror, or p is infinitely great, we have

$$\frac{1}{p} = 0, \quad \text{and therefore} \quad q = \frac{r}{2}.$$

The point O' is in that case at F, half-way between M and C, a result which we have already obtained in § 860.

When p is less than $\frac{r}{2}$, a negative value of q is obtained, showing that in that case the image is formed *back of the mirror* or is *virtual*.

865. Illustrations. If a vase mounted on an open box in which a bouquet of flowers brightly illuminated is hung upside
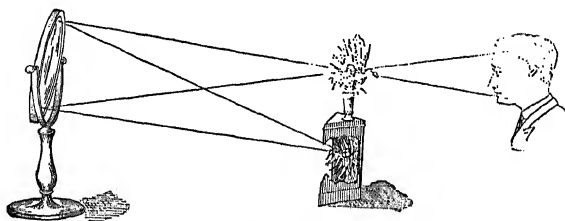
FIG. 503

down, is set in front of a concave mirror at the distance of its center of curvature, and if the mirror is properly inclined, a real image of the bouquet will be formed exactly over the vase, so that to an observer looking over the vase into the mirror the vase appears to hold the flowers. Here object and image are of the same size since equally distant from the mirror.

Standing back of the center of curvature of a concave mirror and looking into it, an inverted and diminished reflection of the face is seen; if, however, the face is held within less than the focal distance MF, the image is virtual, erect, and enlarged, and we have a magnifying mirror.

866. Convex Mirror. In case of a spherical *convex* mirror, the formula obtained in § 864 applies if the radius of curvature is taken negative. Thus for convex mirrors

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{r}$$

expresses the relation between the distances of image and object from the mirror, and its radius of curvature r.

It will be observed that whenever p is positive, it will give a negative value of q, indicating that *wherever the object may be placed in front of the mirror, the image will be formed behind it; that is, it will be virtual.*

The image may be constructed in size and position as before, by tracing from a point P in the object two rays, one which is
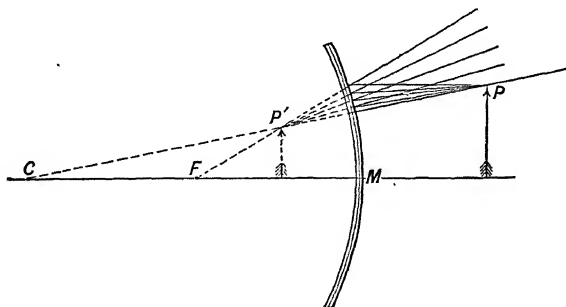
Fig. 504

parallel to the axis and therefore after reflection is directed away from the principal focus F, and another which is directed toward C and meets the mirror perpendicularly so that it is reflected back along the same line. These two rays diverge after reflection and if produced intersect at P' where the virtual image is formed.

The relative sizes of image and object are proportional to

their distances from C. *The image is therefore erect, virtual, and smaller than the object and nearer to the mirror than the object is.*

867. Illustrations. When a polished ball is placed in direct sunlight, the brilliant spot of light seen in the ball is the virtual image of the sun, formed at the principal focus, half-way between the center of the ball and the surface. It is small, for it subtends an angle at the center of the ball equal only to the apparent angular diameter of the sun.

The reflected image of the face seen in a convex mirror is always virtual, erect, and diminished in size.

868. Perfect Mirror. It is useful to consider from the point of view of the wave theory what form a mirror must have to reflect perfectly to a focus all the light that falls on it from a given point. If light waves going out from P as spherical waves
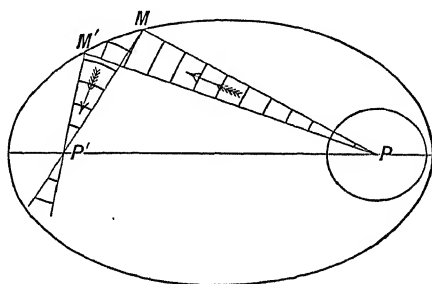
FIG. 505

are to be converged to P', they must after reflection be spherical waves converging toward P'. That is, *all parts of the spherical wave which left P at a given instant must reach P' simultaneously,* and hence it must take light just as long to travel by the path PMP' as by any other path $PM'P'$ where M and M' are points on the mirror, and therefore $PM + MP'$ must be equal to $PM' + M'P'$.

It is known that an ellipsoid of revolution having its foci at P and P' satisfies this condition, hence a perfect mirror should be a portion of the surface of such an ellipsoid. But even then it would be perfect only for light coming from one of its foci. Light from any other point would not be perfectly converged to a single point focus.

869. Aberration. When light from a point in the object does not converge to a point in the image, there is said to be *aberration*. The nature of the aberration in case of a spherical mirror is well shown by reflecting a beam of parallel rays, or light from a distant object such as the sun, in a concave cylindrical mirror mounted over a sheet of white paper, as shown in fig-
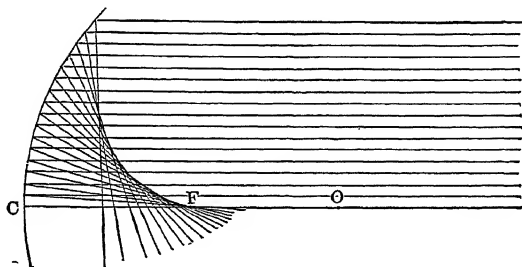
FIG. 506.   Caustic curve

ure 506. It will be observed that only the central rays are reflected through the focus F, those striking the mirror near the edge cross the axis decidedly to the left of F. The curve to which all the reflected rays are tangent is called a *caustic*, and its cusp at F is the ordinary focus of the light reflected from the central part of the mirror.

To avoid excessive aberration the diameter of spherical mirrors is ordinarily small compared with their radius of curvature.

870. Parabolic Mirrors. When it is desired to take a beam of light of large angle and reflect it all in one direction, as in a searchlight, a parabolic mirror is used. For it is a property of the parabola that a line joining any point P with its focus F and a line through P parallel to the axis make equal angles with a tangent at P, and hence a ray of light parallel to the axis will be reflected to F and conversely all rays from F that meet the surface will be reflected parallel to the axis.

FIG. 507.   Parabolic mirror

In searchlights and in locomotive headlights the source of light is not a point but a luminous surface, and an image of this source is formed by the mirror. The size of the image formed is to the size of the source as its distance from the mirror is to the focal length FM. Therefore to illuminate a large region in front of the reflector the source of light should be large and the focal length of the mirror small. While to obtain a very intense beam illuminating only a small patch at a great distance the source should be small and intense and the focal length of the mirror large.

PROBLEMS

1. How high must a plane vertical mirror be in order that an observer 6 ft. in height standing in front of it may just see his whole figure?

2. What sort of mirror must be used and how placed that a ruler in front of the mirror and its image may form two sides of an equilateral triangle?

3. Make a construction showing the size and position of the image formed by a concave mirror having a radius of curvature of 3 in., of an object $\frac{1}{2}$ in. long placed 4 in. in front of the mirror. Make full sized drawing.

4. Make a construction showing the size and position of the image formed by a convex mirror, the object being 4 in. in front of the mirror. Use the same radius of curvature and size of object as in the last problem.

5. A candle is placed 3 ft. in front of a concave mirror having a focal length of $1\frac{1}{2}$ ft.; where is the image, and how large?

6. Where must an arc light be placed in front of a mirror having a radius of curvature of 6 ft. in order that its image may be focused on a screen 20 ft. from the mirror?

7. If a light is placed 2 ft. in front of a concave mirror having a radius of curvature of 6 ft., where will its image be, and how large?

8. How far must a man stand from a concave mirror having a focal length of 2 ft. in order that he may see an erect image of his face just twice its natural size?

9. Which would make the hotter image of the sun, a mirror with a focal length of 6 in. or one with 2 ft. focal length, supposing both to be of the same diameter? Why?

10. How big is the bright image formed when sunlight is reflected by a polished sphere 10 cms. in diameter and where is the image situated? Take the distance of the sun as approximately 110 times its diameter.

11. What sort of mirror must be used and what must be its focal length, in order that it may form an erect image $\frac{2}{3}$ as large as an object placed 2 ft. in front of it? What kind would give an inverted image all other conditions being the same?

REFRACTION

871. Refraction. When a beam of light passes obliquely from one medium into another, it is usually bent at the surface separating the two. This is known as *refraction*. It may be conveniently studied by the aid of the apparatus shown in figure 508.

This consists of a circular glass vessel with flat sides and half-full of water into which a narrow beam of sunlight is directed in a darkened room. If smoke is blown into the space above the water and if the water is very slightly soapy or colored with fluorescein, the path of the beam may be distinctly traced both in the air and water. It is then observed that when the beam is sent vertically downward it is not bent, but when it is inclined it is sharply bent downward at the surface, and the bending is greater the more obliquely the beam meets the surface.

Fig. 508. Refraction of light

The bending also takes place when light passes from water to air, as in case of the coin in the dish shown in figure 509. The coin C is out of sight of the observer's eye so long as the dish is empty, but on filling the dish with water, light coming from the coin is bent into the direction OE and comes to the eye as if from C', and the coin seems lifted into view.

Fig. 509. Coin in dish

In the same way, because of refraction, an oar appears bent upward where it enters the water; and a tank of water, to one looking down into it, looks shallower than it really is, and the more obliquely the bottom is seen the shallower the tank appears.

872. Law of Refraction. The exact law of refraction was discovered by the Dutch physicist, Snell, about 1620, and may be thus stated:

When light passes from one isotropic medium into another, the ratio of the sine of the angle of incidence to the sine of the angle

of refraction is constant for light of any given wave length, whatever may be the inclination of the incident beam, and the incident, reflected, and refracted rays are all in the same plane, called the plane of incidence, which is normal to the surface.

Thus in figure 510, AD and CE are proportional to the sines of the angles i and r, respectively, and the law states that
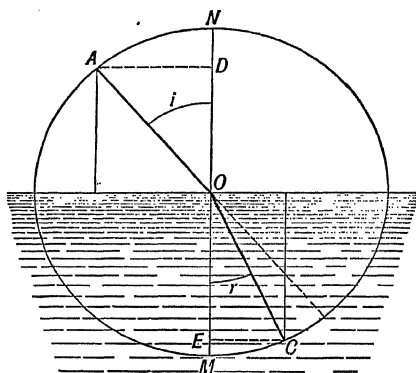
Fig. 510

whatever may be the direction of the incident ray AO, the refracted ray OC will be so inclined that AD will be to CE in a constant ratio which depends on the nature of the two media and on the kind of light. If the upper medium is air and the lower water, AD is very nearly $\frac{4}{3}$ of CE for yellow light, while in case of air and crown glass the ratio of AD to CE is more nearly $\frac{3}{2}$ for the same kind of light.

873. Index of Refraction. This constant ratio of the sine of the angle of incidence to the sine of the angle of refraction is called the *relative index of refraction* of the two media concerned, and the more it differs from unity, the greater the bending of the ray in passing from one medium to the other.

The relative index of refraction when light passes from *air* into a substance is commonly called simply *the index of refraction* of the substance.

The *absolute index of refraction* of a substance is that which holds when light passes from vacuum into the substance; it differs from the ordinary index by only about one part in 3500. It may be determined by multiplying the index from air into the substance by the absolute index of refraction of air, which is 1.000292 at standard conditions.

*Indices of Refraction of Some Common Substances
for Sodium Light*

Glass, very dense flint........................ 1.71
Glass, light crown 1.51
Rock salt 1.54
Diamond 2.47
Water....................................... 1.33
Alcohol 1.36
Carbon bisulphide 1.64
Air 1.000292

874. Cause of Refraction. The velocity of light in water was measured by Foucault and found to be about $\frac{3}{4}$ that in air; and later Michelson found the velocity of light in bisulphide of carbon to be still less than in water. In each case it was found that *the ratio of the velocity of light in air to that in the substance was equal to the index of refraction of the substance.* Let us now inquire whether the assumption that a beam of light consists of a train of waves which experience a change of velocity in passing from one medium into another will account for the above result, and also whether it affords a satisfactory explanation of the law of refraction as established by experiment. In the following paragraphs we shall trace the consequences of this assumption.

875. Perpendicular Incidence: No Change in Direction. When a beam of light in air is perpendicular to the surface of another substance, as water, in which its velocity is less, the wave fronts are parallel to the surface AB, and consequently all parts of a given wave front meet the surface AB at the same instant, and advancing into the lower medium with the same velocity everywhere, the wave

Fig. 511. Perpendicular incidence of waves

front in the lower medium must remain parallel to the surface AB, and the ray direction remains unchanged.

The wave length, however, in the lower medium must be less than in air in the same ratio as the velocity of light in the substance is less

than its velocity in air. For it must advance one wave length in the substance in the same time that it advances one wave length in air, since just as many waves per second enter the lower medium as leave the air.

876. Oblique Incidence: Change in Velocity and Direction. If the incident beam falls obliquely on the refracting surface then the change in velocity in passing from one medium to the other causes a bending of the ray or change in direction, as shown in figure 512.

For, let the heavy lines represent wave fronts one wave length apart, advancing in the direction of the arrows, and let the second medium be one, such as glass, in which the velocity is less than in air. As soon as the edge of the wave enters the glass at A it is retarded, while that part which is still in air continues to advance with the same velocity as before. Consequently the direction of the wave front is changed into the position DC.

FIG. 512. Refraction of oblique waves

Now, BC is the distance that a wave travels in the upper medium in the same time that it travels a distance AD in the lower medium; therefore

$$BC : AD :: V : v$$

where V is the velocity of light in the upper medium and v its velocity in the lower one.

Let i be the angle of incidence BCN or BAC and let r be the angle of refraction ECM or ACD, then

$$BC = AC . \sin i$$
$$AD = AC . \sin r$$

hence, dividing, we find

$$\frac{BC}{AD} \quad \frac{\sin i}{\sin r}$$

but

$$\frac{BC}{AD} = \frac{V}{v};$$

therefore

$$\frac{sin \cdot}{sin \cdot} = \frac{V}{v} = n,$$

the refractive index.

From this it appears that *the ratio of the sine of the angle of incidence to the sine of the angle of refraction is the same as the ratio of the velocities of light in the two media, and must, therefore, be constant for all angles of incidence.*

877. Adequacy of the Wave Theory. The above interesting result is in exact agreement with the law of refraction as discovered by .Snell, and it also leads to the conclusion that *the relative index of refraction of two media is simply the ratio of the velocities. of light in those media,* a conclusion substantiated by the measurements of the velocity of light in water and in bisulphide of carbon by Foucault and Michelson.

We thus find that the wave theory leads to a simple and natural explanation of the facts known about refraction, a result which must strengthen our conviction of the essential soundness of the theory.

Also the index of refraction of a substance takes on a new interest when we think of its physical significance as the ratio of the velocity of light in air or vacuum to that in the substance.

878. Total Reflection. When light passes from one medium into another in which the velocity of light is *greater,* as when it passes from water or glass into air, the refracted ray is bent *away* from the normal. Thus a ray of light coming up from below and meeting the surface of water on the under side, as shown by *A O* in the first diagram of figure 513, is in general partly refracted, and bent away from the normal in the direction *OB* and partly reflected along *OD.* But as the direction of *A O* is changed and made more oblique, *OB* is bent away more strongly until when *A O* takes the direction shown in the middle diagram of the figure, the refracted ray *OB* emerges at an angle of 90° and grazes along the surface. The angle *A ON* in this case is called *the critical angle.* When the angle of incidence is greater than the

critical angle, as shown in the third diagram, none of the light is refracted but the beam is totally reflected along OD, as if the surface of the liquid were a polished metal mirror, for there is no
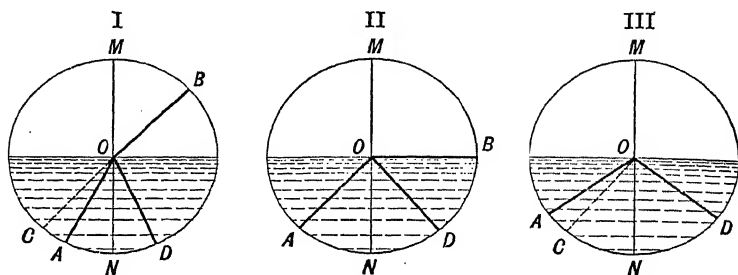
FIG. 513

corresponding direction in which it can emerge into the upper medium.

879. Critical Angle. The angle AON in the middle diagram above, beyond which refraction cannot take place, is called the *critical angle*. From the law of refraction,

$$\frac{sin\ MOB}{sin\ AON} = n,$$

but $sin\ MOB = 1$ since MOB is a right angle, therefore

$$sin\ AON = \frac{1}{n}$$

or the *sine of the critical angle is equal to the reciprocal of the index of refraction.*

880. Illustration of Total Reflection. If a tumbler full of water and having smooth sides is held in the hand, on looking down obliquely into it the sides are seen as polished, mirror-like surfaces reflecting objects under the glass but the fingers holding the glass cannot be seen through the surface if it is dry, as in that case light coming up from below is totally reflected at the side. If the fingers are moist they will be seen only at the spots where they press against the glass.

A right-angled glass prism having all its sides polished may be used as a mirror to turn a beam of light through 90° if the light falls upon it as shown in the figure, for in that case it meets the oblique surface *inside* the glass at 45°, which is greater than the critical angle for glass and air. The intensity of the beam reflected in this way is far greater than if reflected from the *outside* of the same surface, for in that case a large amount of light is lost by refraction through the prism.

In figure 515 is shown a right-angled prism used as a reversing prism with a projecting lantern. The beam AA' which on entering the prism is directed downward, on leaving it is sloping upward, so also BB' is changed from an upward inclination on entering the prism to an equal downward slope on emergence.
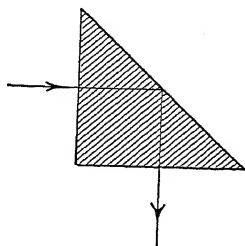
Fig. 514. Total reflecting prism

881. Refraction of Gases. The refracting power of gases is small compared with that of solids or liquids, the change in
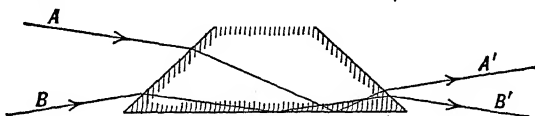
Fig. 515. Reversing prism

velocity when light passes from vacuum into air under ordinary conditions being only about one nine-hundredth part of the change in velocity when it enters water. Yet it is the variations of this small refractive power caused by the fluctuating density in the hot-air currents over a stove that cause the unsteadiness in the appearance of bodies seen through the stream of hot air.

882. Atmospheric Refraction. In consequence of the refraction of the air the apparent angular distances of stars from the zenith are less than their true zenith distances, the rays being refracted just as much as if the atmosphere terminated abruptly in a level surface just above the observing telescope and all above were vacuum, instead of gradually diminishing in density

as it does. The sun or moon when seen near the horizon appears flattened in consequence of the lower edge being more raised by refraction than the upper edge, and when apparently just above the horizon it is really entirely below it.

883. Mirage. When a layer of air next the surface of the earth becomes heated it may become less dense and less refracting than the cooler layers above it, so that the lower edges of light
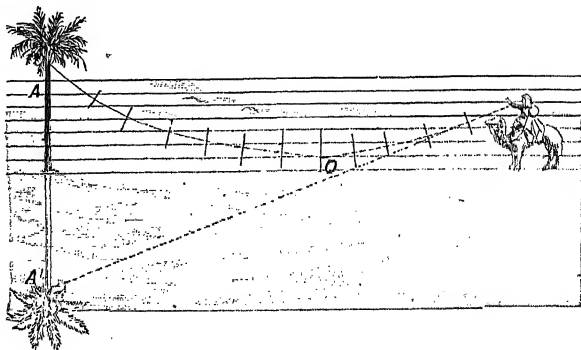
FIG. 516

waves coming from a distant object are less retarded than the upper parts of the waves, and consequently the wave fronts swing around and come upward to the eye, as shown in figure 516. The distant object is thus seen inverted as if reflected in a horizontal mirror. In this way the familiar mirage of the desert may give the impression that objects seen are reflected in a sheet of water.

PRISMS AND LENSES

884. Refraction of Plane Waves by Plate with Parallel Sides. In passing into the plate the beam is bent toward the normal, but since the two sides are parallel the waves within the plate make the same angle with one side as with the other and will therefore be bent as much on emerging from the plate as they were bent on entering, and the emergent beam will therefore be parallel to the entering one, but displaced sidewise by an amount which depends on the thickness of the plate. Light

waves from a *distant* point will therefore enter the eye of an observer in the same direction as if the plate were not there.

If the apparent position of a star shifts on interposing a piece of thick plate glass, even if held obliquely, it is because the sides of the plate are not perfectly parallel.
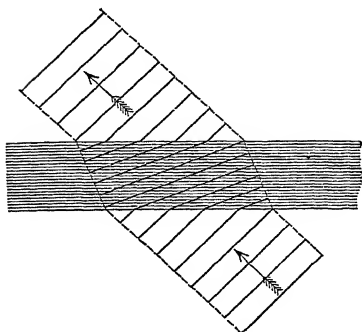
FIG. 517. Refraction through plate with parallel sides
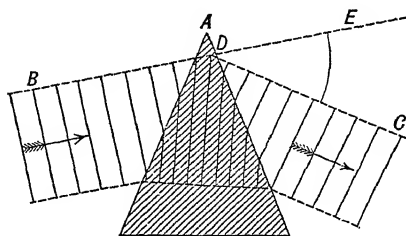
FIG. 518. Refraction through a prism

885. Refraction by a Prism. Plane waves when refracted at a plane surface remain plane, and therefore will continue plane after any number of successive refractions at plane surfaces.

When a substance has two plane refracting surfaces which are inclined to each other it is called a prism, and the angle between the two refracting surfaces is called the angle of the prism.

In figure 518 the edge of the prism at A is supposed to be perpendicular to the plane of the paper, which is the plane of incidence. The beam of light at B enters the prism, is bent aside, and on emergence is again bent, and passes out in the direction shown at C. The total change in direction is represented by the angle CDE, which is called the deviation of the beam.

The beam is bent *toward the thicker part of the prism*, as shown in the figure, when the substance of the prism is more refracting than air, because that part of each wave is most retarded which is farthest from the edge of the prism and has to pass through the greatest thickness of retarding substance.

886. Minimum Deviation. In such a position of the prism as that shown in figure 519, in which the incident beam makes the same angle with the first face of the prism as the emergent beam does with the second, it is found that the deviation angle *CDE* is a *minimum;* turning the prism away from this position in either direction causes the angle *CDE* to increase.

If *n* represents the index of refraction of the substance of the
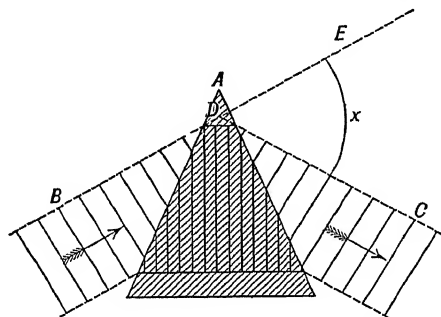
Fig. 519.　Minimum deviation

prism, and if *A* is its angle and *D* the angle of minimum deviation, it may easily be proved that

$$n = \frac{sin \cdot \dfrac{A + D}{}}{sin \cdot}$$

887. Lenses. Lenses are pieces of glass or other transparent substance usually bounded by spherical surfaces, and are used

3　　　　　　　　　　　　　　6

Fig. 520

in forming optical images. The line joining the centers of curvature of the surfaces of a lens is called its *axis.* Different

types of lenses are shown in figure 520. These are distinguished as double convex (1), plano-convex (2), meniscus (3), double concave (4), plano-concave (5), and convexo-concave (6).

In the first three cases light rays parallel to the axis are converged to a point F, called the principal focus. The distance of this point from the lens is called its focal length. Such lenses are called *convergent;* they are thicker in the center than at the
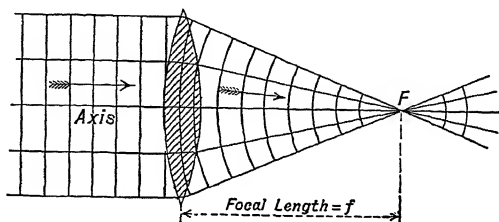
FIG. 521. Convex lens. Focal length $= f$

edges, and consequently plane waves passing through them are more retarded at the middle than at the edges, and become of a concave spherical form converging on F.

The last three forms of lens are thinner at the center than at the edges and are known as *divergent* lenses, for plane waves
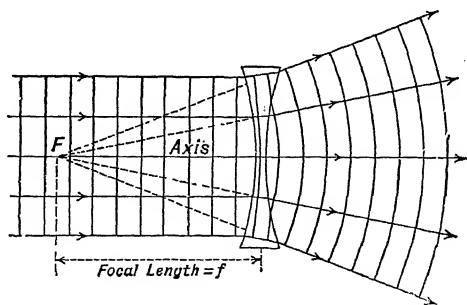
FIG. 522. Concave lens. Focal length $= f$

advancing along the axis of such a lens are more retarded at the edges than at the center and emerge from the lens as spherical waves expanding from a center F. This point from which rays parallel to the axis on one side of the lens appear to diverge on the other side is called the principal focus. In this case it is a *virtual* focus.

Substituting this value of $r + r'$ in (1), we have $i + i' = n(b + d)$, and now substituting this in (2) we find

$$n(b + d) = a + c + b + d$$

or

$$a + c = (n - 1)(b + d). \qquad (3)$$

Now if the lens is thin E and F are practically at the same distance from the axis; call this distance y, and let p be the distance AE, q the distance CF, while R_1 and R_2 represent DE and BF, the radii of curvature of the lens surfaces. Then if the angles $abdc$ are small, each will be equal to the arc that subtends it divided by the corresponding radius, but all the arcs may be considered equal to y, so that

$$a = \frac{y}{p} \qquad b = \frac{y}{R_2} \qquad c = \frac{y}{q} \qquad d = \frac{y}{R_1}$$

Substituting in (3) we have

$$\frac{y}{p} + \frac{y}{q} = (n - 1)\left(\frac{y}{R_1} + \frac{y}{R_2}\right)$$

or finally

$$\frac{1}{p} + \frac{1}{q} = (n - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)$$

which may be written

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

where

$$\frac{1}{f} = (n - 1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)$$

and f is a constant for the lens, depending on its index of refraction and the radii of curvature of its surfaces.

889. Discussion of Formula. Consider first the case of a convergent lens with the source of light an infinitely distant point, in this case f is positive and $p = \infty$; therefore, $\dfrac{1}{p} = 0$ and $q = f$.

Therefore the light converges to a point at a distance f from the lens. This point is the *principal focus* and the distance f is the *focal length* of the lens.

As the point source P is moved along the axis nearer to the lens, the corresponding focus Q or *conjugate focus* moves away

from the lens, so that when $p = q$, each of the points P and Q is at a distance from the lens equal to $2f$.

As P is now moved uniformly toward the lens, the rays on the farther side become more nearly parallel and Q moves off with increasing speed till the distance p is equal to f when q becomes infinite and the rays go out from the lens parallel.

If P is now moved still nearer to the lens, the rays on the
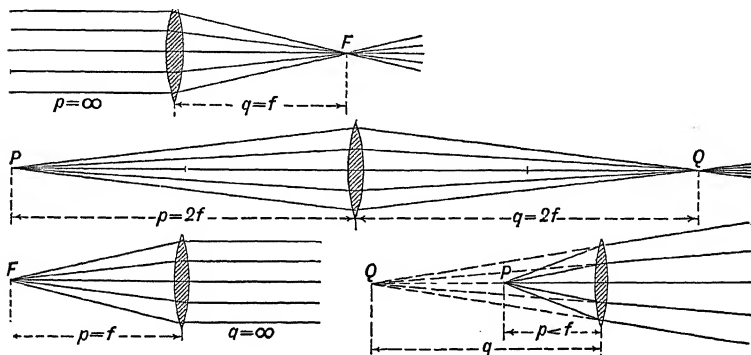
FIG. 526

farther side diverge as if they came from a focus Q on the left of the lens. In this case the formula shows that q will be *negative*, and the focus Q is *virtual*.
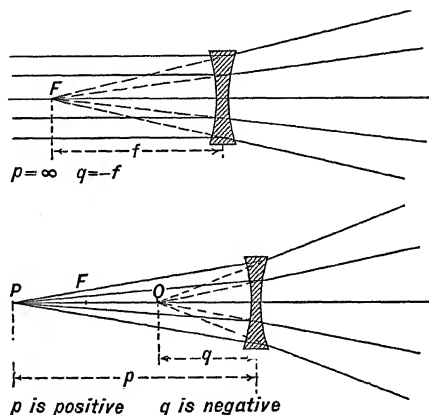
FIG. 527.   Divergent lens

In case of a *divergent* lens, the focal length is *negative*, so that we have

$$\frac{1}{p} + \frac{1}{q} = -\frac{1}{f}.$$

Here if the luminous point P is at an infinite distance, we find $q = -f$, indicating that the rays diverge as if from a point F. This is the *principal focus* and it is virtual.

As P moves from an infinite distance in toward the lens, the *conjugate focus* Q remains virtual and moves from F toward the lens, so that when $p = f$, $q = -\frac{f}{2}$, and when $p = 0$, $q = 0$.

890. Focus for Distant Objects.   It is important to observe that *when the point P is at a great distance from the lens compared with its focal length, the conjugate focus Q is very nearly at the principal focus*, and a great change in the position of P will cause only a slight change in Q.   Thus when $p =$ ten times f, $q = \frac{10}{9}f$, while if p is 100 times f, $q = \frac{100}{99}f$.

It is for this reason that in a photographic camera the focus for all *distant* objects is practically the same.

891. Rule for Use of Formula.   In using the formula

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

in the solution of lens problems care must be taken as to the *signs* of the various terms.

The focal length f is always *positive in case of a convergent lens and negative in case of a divergent lens.*

When the rays from the point P diverge toward the lens the sign of p is positive; if, however, the rays meet the lens as they are converging toward P then p must be taken *negative*.

So also *if rays leaving the lens converge toward a real focus the distance q of that focus from the lens is positive,* while if the rays after passing the lens diverge from a *virtual focus* the distance q is *negative* and is measured back of the lens.

Notice that we always consider a narrow pencil of rays *originating in a single point* in the object; and that p is determined by the rays of such a pencil as they *approach* the lens, while q relates to the rays *leaving* the lens.

The rule of signs may be illustrated by the case shown in figure 528. A convergent lens of 4 in. focus is placed 4 in. from a divergent lens of the same focal length; if an object P is placed 6 in. from the convergent lens, it is required to find the position of the image formed by the combination.

Rays from P diverge toward the lens L, therefore the distance 6 is positive, and as the focal length of that lens is also positive we have

$$\frac{1}{6} + \frac{1}{q} = \frac{1}{4}$$

which gives $q = +12$.

Hence the rays after passing the first lens converge toward a real focus at Q, 12 in. to the right of L.
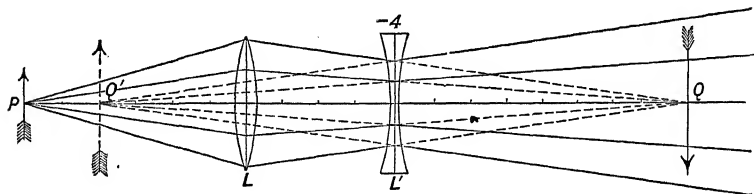
FIG. 528

But in case of the second lens L' the rays approaching it are converging toward a point Q which is 8 in. beyond the lens; hence in this case $p = -8$, and as the lens is divergent $f = -4$; hence we have

$$-\frac{1}{8} + \frac{1}{q} = -\frac{1}{4}$$

which gives $q = -8$. The rays will therefore emerge as if coming from a point Q', 8 in. back of the lens L', and the final image is *virtual*.

892. Images by Lenses. The most important use of lenses is in the formation of optical images. Let P (Fig. 529) be a point on the arrow that lies on the axis of the lens; light from P will be converged at Q, its conjugate focus. So also for any other point P' in the object there is a conjugate focus Q' which must lie on the *secondary axis* or straight line $P'OQ'$ through the *center* of the lens, for at the center the opposite faces of the lens are

parallel and hence a ray passing through the center is not changed in direction. If P and P' are in the same plane at right angles to the axis of the lens, Q' will not be in the parallel plane through Q, but will be somewhat nearer the lens, making the image curved.

If a white screen is placed at Q the light falling upon each point of it comes from the corresponding *conjugate* point on the
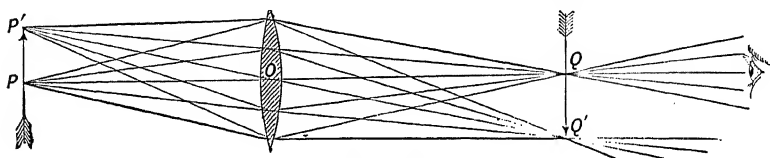
Fig. 529

other side of the lens and a picture or real image is therefore formed on the screen just as though the light had come through a pin hole at O, but more brilliant. This is the principle of the photographic camera.

The image formed by a lens may be seen directly by the eye instead of being received on a screen; for the eye may be placed to the right of Q at a distance from it of about 10 in., or the distance of normal distinct vision, and looking toward Q light will enter the eye from Q just as it would have come from an object placed at that point, and accordingly the inverted image of the arrow will be seen.
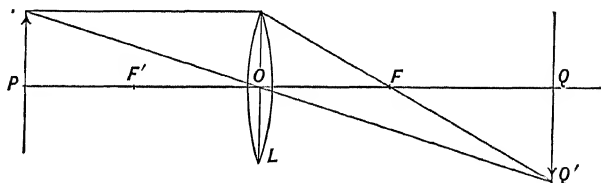
Fig. 530

893. Construction of Images. A simple geometrical construction will give the size and position of the image in any case, for *it is only necessary to trace two rays from any point in the object to find by their intersection the position of the corresponding point of the image.* Suppose it is required to find the size and position of the image of the arrow at P formed by the convergent lens L whose principal foci are at F and F' (Fig. 530). Since the

size as well as the position of the image is desired, we will choose a point P' not on the axis of the lens and trace two rays. One parallel to the axis must after refraction by the lens pass through the principal focus F. Another ray through the center of the lens O is undeviated, and where these two meet at Q' is the image of the point P'. Since the image of P must be formed at Q, the
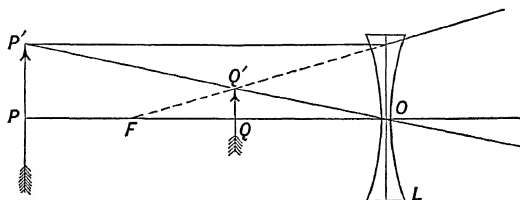
FIG. 531

length of the object PP' is to the length of the image QQ' in the same proportion as their distances from the lens.

By a similar construction the size and position of the image formed by a divergent lens may be found, as in figure 531, where F is the principal focus of the lens L. A ray from P' approaching the lens parallel to the axis is refracted up as if it came from the principal focus F, while the second ray through the center is un-

FIG. 532

deviated. The two rays after emerging from the lens diverge as if from the point Q', which is therefore the *virtual* image of the point P'.
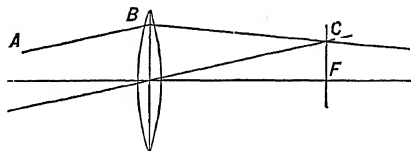
The relative size of object and image is, as before, the ratio of their distances from the lens.

894. Thick Lenses. It was shown by Gauss that for thick lenses or for a combination of lenses there are two *principal planes* perpendicular to the axis such that rays on one side of the lens which are directed toward any point such as A (Fig. 533) in one plane will on emergence be directed as if from the opposite point A' on the other plane, and the points HH' (known as principal points) where these planes meet the axis have the additional property that rays directed toward H emerge in a *parallel* direction from H'. The principal focus F on one side is just as far from H as the other focus F'

is from H'. This distance from either principal focus to the corresponding principal plane is called the *focal length* of the lens. If p and q are the distances of object and image, also measured from H and H' respectively, the simple formula

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$$

holds just as in case of thin lenses *for all pencils of light that are but slightly oblique to the axis.*

The graphic construction of images, using the principal planes, is precisely similar to that explained above in § 893 for thin lenses, except that $H'Q'$ is to
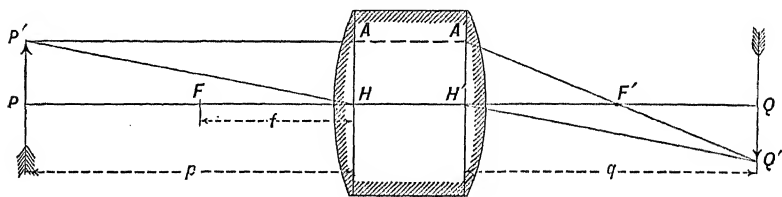
Fig. 533

be drawn *parallel* to $P'H$ instead of simply prolonging $P'H$. If the planes AH and $A'H'$ are made to coincide, the construction is that for a thin lens.

895. Defects of Images Formed by Lenses. Besides the curvature of the image which is noticed when the object has large angular dimensions
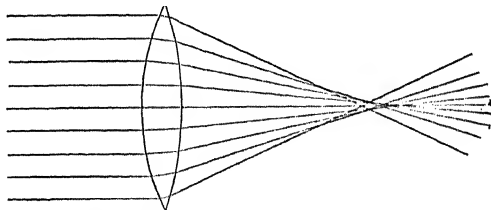
Fig. 534. Spherical aberration

as seen from the lens, there are other defects in the images formed by lenses which tax the skill of the optician to overcome.

Almost all lenses have spherical surfaces and are subject to a defect called *spherical aberration.* One kind of spherical aberration is that rays from a given point which are refracted on different portions of the lens do not meet accurately at a single focus, the rays refracted by the outer portions of the lens coming to a focus nearer the lens than those passing through its central region as indicated in figure 534. Another kind of spherical aberration is that a pencil of rays from a point, passing obliquely through the lens, becomes astigmatic and converges through two focal lines instead of coming to a single point. These defects cause a lack of clearness and sharpness in

the images formed. They are most serious when the diameter of the lens is a large fraction of its focal length.

The colors observed at the edges of images formed by lenses are due to the fact that ordinary lenses refract blue light more strongly than red light. This defect, known as chromatic aberration, will be discussed later (§ 902).

PROBLEMS

1. How deep is a tank of water which appears to be 4 ft. deep to a person looking vertically down into it?

2. An incandescent lamp is placed 6 ft. below the surface of a pond. Show why only a fractional part of the light can escape directly from the water.

3. If a beam of light has 50,000 light waves to the inch in air, how many to the inch will there be after it has entered water?

4. Find the velocity of light in water if the critical angle at the surface between water and air is 48° 30'.

5. When the index of refraction of water is 1.33 and that of carbon bisulphide is 1.67, what is the critical angle between water and carbon bisulphide?

6. The object-glass of the Yerkes telescope is a convergent lens 40 in. in diameter and having a focal length of 62 ft. What is the size of the sun's image formed by it? What effect has the size of the lens on the size of the image? Assume that the sun's disc subtends an angle of 31' 14.1''.

7. An incandescent lamp is 30 cms. from a convergent lens of 10 cms. focal length. Find the position and relative size of the image; is it real or virtual?

8. A candle is placed 1 meter from a divergent lens having a focal length of 1 meter. Where is the image formed and what is its size? Make a construction illustrating the case.

9. A lamp and a screen are 10 ft. apart. Where must a convergent lens of 2 ft. focal length be placed so as to form an image of the lamp on the screen? Show that there are two solutions and find the relative size of the image in each case.

10. A beam of sunlight falls on a divergent lens of focal length 10 in.; 20 in. beyond this lens is placed a convergent lens of 15 in. focal length. Find where a screen should be placed to receive the final image of the sun.

11. A convergent lens, focal length 10 in., is placed 12 in. from a gas flame; then 36 in. beyond the first lens is placed a divergent lens of focal length 16 in. Find the position and size of the final image; is it real or virtual?

12. A certain lens when placed 10 cms. from an object, forms a virtual image 5 times as large as the object. What kind of lens is used and what is its focal length?

13. What must be the focal length of spectacle lenses so that a man who can see distinctly objects 2 meters distant without the glasses can read print at 40 cms. distance with them, and what kind of lenses must be used?

NOTE: The strength of spectacle lenses is expressed in diopters and is the reciprocal of the focal length expressed in meters.

14. A person, who without glasses cannot see distinctly objects more than 12 cms. from the eye, wishes glasses to enable him to see clearly distant objects. What must be the kind used and their focal length and strength in diopters?

DISPERSION

896. **Dispersion of Light by a Prism.** When a narrow beam of sunlight passes through a prism, the light is not only bent aside or *deviated*, it is also *dispersed* or spread out into a colored band called *the spectrum*.

Sir Isaac Newton placed a second prism (Fig. 535) in the spectrum so that light of only one color might fall on it. This
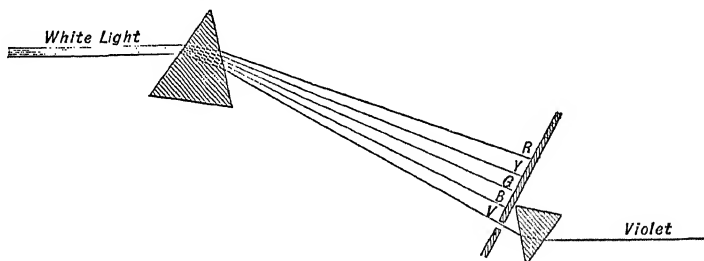
FIG. 535. Newton's experiment

light was refracted on passing through the second prism, but there was no further change in color, showing that the prism itself did not produce the different colors, but simply separated the various kinds of light already present in the beam of sunlight. The separation is effected because the various colored lights are differently refracted by the prism, the red being refracted least and the violet most.

897. Cause of Dispersion. Since the bending of the rays by a prism depends only on the angle of the prism and the index of refraction of the substance of which it is made, it follows that *the index of refraction of the prism must be different for each kind of light in the spectrum*, being least for the red which is

least refracted and greatest for the violet which is most strongly refracted.

Of course *the interpretation of this fact is that red light must pass through the substance of the prism with greater velocity than violet light.* It will be shown later that the physical difference between one kind of light and another lies in their wave lengths. These vary from one end of the spectrum to the other, the longest waves being at the red end of the spectrum while the shortest are at the violet end.

It appears, therefore, that shorter waves of light are more retarded in passing through glass than longer ones.

898 **Dispersive Power.** When two prisms of different substances have such angles that each produces the same deviation
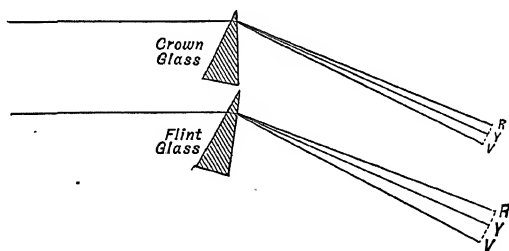
FIG. 536. Prisms with different dispersive powers

for yellow light, or light in the middle of the spectrum, the angular widths of the two spectra produced will usually not be the same, but are proportional to what are called the *dispersive powers* of the substances. The *dispersive powers* of some substances are as follows:

Water 0.042
Carbon bisulphide 0.145
Crown glass 0.043
Flint glass 0.061

Thus for an equal bending of the mean rays, carbon bisulphide will produce a spectrum $3\frac{1}{3}$ times as long as that produced by crown glass and $2\frac{1}{3}$ times as long as one formed by flint glass.

899. **Calculation of Dispersive Power.** We have seen that the index of refraction of a substance depends on the *kind* of light. The following table gives three indices of refraction for each of four substances. The indices given in the first column are for light near the extreme red end of the spectrum, those in the second are for the yellow sodium light, while those in the third are for light near the violet end of the spectrum. These points in the spectrum correspond to three dark lines in the sun spectrum designated A, D, and H by Fraunhofer (§ 932).

<center>INDICES OF REFRACTION</center>

SUBSTANCE	n_A	n_D	n_H	$n_D - 1$	$n_H - n_A$
Water at 16° C.............	1.330	1.334	1.344	0.334	0.014
Carbon bisulphide at 10° C...	1.616	1.635	1.708	0.635	0.092
A kind of crown glass........	1.528	1.534	1.551	0.534	0.023
A kind of flint glass	1.578	1.587	1.614	0.587	0.036

The next to the last column in the above table shows *the relative deviations of yellow sodium light caused by prisms of the different substances all having the same small angle.* Thus it appears that a prism of carbon bisulphide will cause nearly twice as great a deviation as a prism of water of the same angle, if the angles of the prism are small.

In the last column are given the *differences* between the indices of refraction for red and violet lights, which represent *the relative angular widths of the spectra produced by prisms of the various substances having the same small angle.* The spectrum formed by a thin prism of carbon bisulphide is therefore about $6\frac{1}{2}$ times as long as that formed by a similar prism of water.

To obtain the relative *dispersive powers* given in the previous paragraph the figures in the last column must be divided by those in the next to the last.
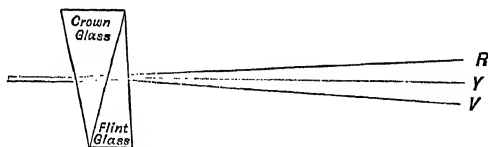
<center>FIG. 537. Dispersion without bending the mean ray</center>

900. Direct-vision Prism. In consequence of the fact that the dispersive powers of substances differ it is possible to so combine two prisms of different substances as to produce dispersion without deviation of the mean ray, or to produce deviation without dispersion.

For example, if a prism of crown glass and one of flint glass

are taken whose angles are small, and in the ratio 0.587 to 0.534, respectively, they will each deviate the D line of the spectrum by the same amount (see table § 899), but the spectrum formed by the flint prism will be longer than that formed by the crown in the ratio of 0.061 to 0.043 (§ 898). If, therefore, the two are
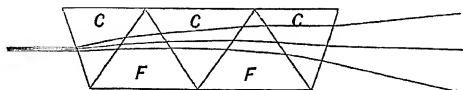
FIG. 538. Amici prism

placed with their edges oppositely directed, as shown in figure 537, the deviation of one will be balanced by that of the other for the D line or yellow light of the spectrum, but as the dispersive power of the flint is greater than that of the crown there will still be a spectrum formed with the violet toward the base of the flint and the red toward its edge.

Such a combination is known as a direct-vision prism. By using two prisms of flint combined with three of crown of suitable angles, as shown in figure 538, a very large dispersion may be produced with no deviation of the middle part of the spectrum.

901. Achromatism. It is, however, of much more practical importance to produce *deviation without dispersion*. To obtain this result two prisms of crown and flint glass may be combined whose angles are in the ratio 0.036 to 0.023 or inversely as the ratio of the angular width of their spectra given in the last column of the table in § 899. Two such prisms will give
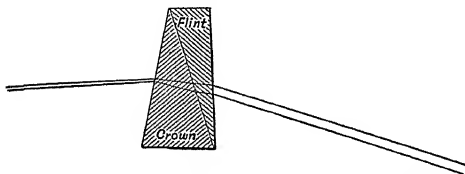
FIG. 539. Ray bent without dispersion

spectra of the same angular width, but the deviation by the crown-glass prism will be greater than that by the flint. If the two are now placed together so as to act oppositely, as shown in figure 539, the beam of light will be deviated toward the base of the crown-glass prism, but there will be no dispersion,

since in this respect the two balance each other. Such a prism is called *achromatic*.

(902) **Achromatic Lens.** If sunlight passes through an ordinary convergent lens made of a single piece of glass, it may easily be shown, by interposing successively a red glass and a blue glass, that the focus for red light is at a greater distance from the lens than that for blue light. For every little portion of the lens acts as a prism bending light toward the axis and at the same time dispersing it. With such a lens there is no point at which a sharp image of an object will be formed by ordinary white light. All points in the image will be blurred
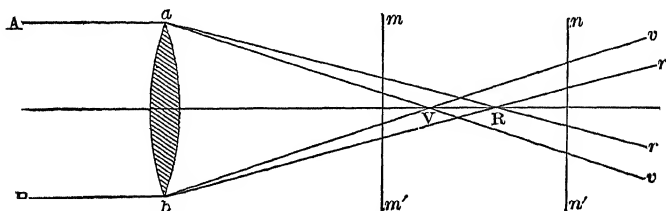
FIG. 540. Different foci for violet and red rays

and all lines of separation between light and dark portions of the image will be colored.

This serious defect may be remedied by combining a convergent lens of crown glass with a divergent lens of flint, as shown in figure 541, to form a *convergent achromatic lens;* or if the crown-glass lens is divergent and the flint convergent, a *divergent* achromatic lens may be formed. In either case the curvatures must be so chosen that each little portion of the combination through which a ray passes will act like an achromatic prism as explained in the previous paragraph.

Convergent   Divergent

FIG. 541. Achromatic lenses

In case the two component lenses are in contact* and their curvatures are not great, *achromatism will be produced if the*

---

* It is common in small lenses to cement together the two lenses of an achromatic combination with Canada balsam in order to prevent loss of light by reflection from the inner surfaces.

focal length of the crown-glass lens is to that of the flint in the same proportion as their dispersive powers (§ 898).

If the ratio of the dispersive powers of two kinds of glass, such as flint and crown, were the same in all parts of the spectrum, a perfectly achromatic lens could be formed in this way, but unfortunately this is not the case with ordinary glasses, so that the image formed even by an achromatic lens shows some residual color.* Lenses intended for visual observation are made so as to bring as nearly as possible to one focus the orange, yellow, and green rays which form the brightest part of the spectrum; while for photographic purposes the lens must be achromatic for the violet and ultraviolet rays which are most strongly *actinic* or active on a photographic film.

903. Ordinary and Anomalous Dispersion. In most substances the shorter the wave length the more strongly the light is retarded or refracted.
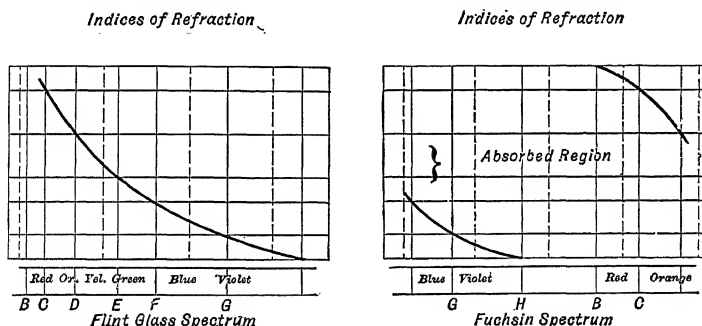
FIG. 542. Dispersion curves

This is called ordinary or *normal dispersion*. But many substances exhibit what is called *anomalous dispersion*, especially such as absorb very strongly light of some particular wave lengths while transmitting comparatively freely those waves which are slightly longer or shorter than the ones absorbed. Figure 542 shows the normal dispersion curve of flint glass in which the index of refraction increases as wave lengths decrease. But the curve for fuchsin shows that when sunlight is dispersed by a prism of this substance the middle part of the spectrum including green is wanting, being completely absorbed, while red and orange, which in normal dispersion

* Certain kinds of glass are now made at Jena from which compound lenses are made which give images almost free from residual color.

would be less refracted than blue, are so much more strongly refracted that they are beyond even the extreme violet. This substance thus illustrates the general law first stated by Kundt, that *the effect of selective absorption is to increase the velocity of those light waves which are somewhat shorter than the ones most strongly absorbed and to retard light of greater wave length*.

For an interesting discussion of anomalous dispersion, see Edser, *Light for Students*.

904. Rainbow. When parallel rays of light fall on a spherical drop of water, some of the light is refracted into the drop, suffers reflection at the opposite surface, and is then refracted out again. The direction in which it
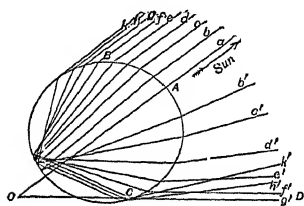
FIG. 543. Refraction in rain drop

finally comes from the drop depends on the point where it entered. A ray *a* falling on the center at *A* is reflected back on its path, while *b* just above the center will be refracted as shown at *b'*. Rays meeting the drop farther from *A* are still further inclined downward on emergence, until we come to a group of rays, *efg*, meeting the surfaces at *B*, which have the maximum downward direction as shown at *e'f'g'*; rays beyond *B* are again turned upward as shown at *h'* and *k'*. The pencil of rays refracted at *B* do not scatter on leaving the drop, but emerge as a nearly parallel beam in the direction *CD*. If the eye is so placed as to receive this beam the drop will appear very bright, while if the eye is above the line *CD* the drop will appear but faintly illuminated in consequence of the scattering of the emergent rays, but if the point of sight is *below* the line *CD*, the drop will appear dark, for no light at all is sent in such a direction.

The angle *AOD* between the bright *beam CD* and the original direction *AO* depends on the index of refraction of the drop, and varies with the wave length, being about 42° for red light and 40° for violet.

In consequence of this, drops of dew seen in bright sunlight at the proper angle may appear as brilliant jewels colored red, yellow, green, or blue. In such a case it will be found that a slight change in position of the eye may cause a red drop to change to green or blue.

The formation of the primary rainbow may now be readily understood from figure 544. Let *ABCD* represent drops in the air all illuminated by the sun's rays from *S*. From each drop there will emerge a bright parallel pencil of red light in the direction *R* making an angle of 42° with *S*, and a bright violet pencil *V* making an angle of 40° with *S*, the intermediate colors of the spectrum being between these extremes. An observer, therefore, whose eye is at *E*, will receive red light from the drop *B* and violet from the drop *C* and intermediate colors from drops between them. If the diagram is now conceived to be rotated about the line *NES*, joining the eye and the sun, it is clear that every drop on the circle whose radius is *BN*, which is described by the motion of the drop *B*, will send red light to the eye, while all drops on the circle described by the motion of *C* will send violet light to the

eye. In this way a colored circular band will be seen whose angular radius is between 40° and 42°.

It will be observed also that the drop A does not send any light at all to the eye at E, while scattered rays of all colors come from the drop D. The region, therefore, above the primary rainbow appears dark, while that within
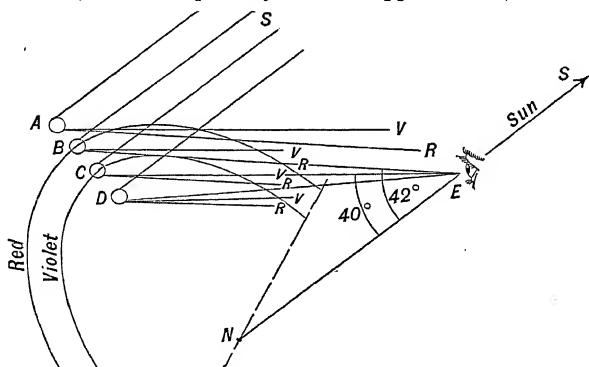
Fig. 544. Primary rainbow

it is bright, and the red of the bow is nearly pure, while the violet is mixed with scattered rays of other colors and fades out into white.

905. Secondary Rainbow. For those rays that suffer *two* reflections inside a rain drop there is also a certain direction in which the emergent
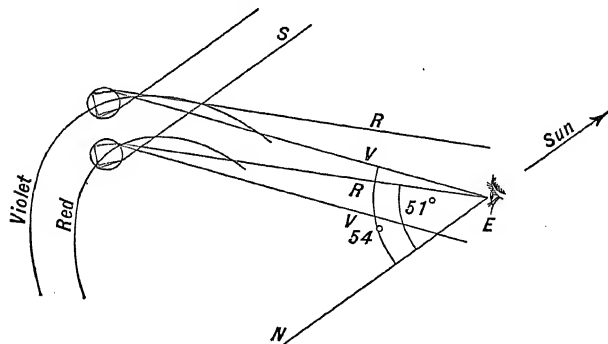
Fig. 545. Secondary rainbow

rays are parallel, and therefore the light in that direction is particularly intense.

A colored bow will, therefore, be produced as shown in figure 545, the angular radius of the red being 51° while that of the violet is 54°. The sky within this bow and between it and the primary bow will be dark while outside of it beyond the violet the sky will be bright with scattered light.

906. Supernumerary Bows. The bows caused by more than two internal reflections cannot be seen. A second and even a third band of red may, however, be occasionally seen in the violet region of the primary bow. These are called *supernumerary* bows and are *diffraction* phenomena (§ 963). Their explanation is given in more advanced treatises, such as *Preston's Theory of Light.*

Optical Instruments

907. Optical Instruments. There are two general classes of optical instruments, those which form a real image on a screen, as in case of the photographic camera and projection lantern (magic lantern), and those intended for direct eye observation in which the image formed is virtual. To the latter class belong the magnifying glass, microscope, and telescope.

To obtain a clear conception of the action of an optical apparatus it is desirable to study the effect of the instrument upon two pencils of light, starting from different points in the object and traced through to the corresponding points in the image. One pencil should be as oblique as can pass through the instrument.

908. Photographic Camera. In the simplest form of photographic camera a single convergent lens forms a real image of a distant object on the sensitive plate. A diaphragm placed close to the lens limits the size of the pencil of light. The *quickness* of a photographic lens or the brightness of the image will be proportional to A, the area of the diaphragm opening, and inversely proportional to the area of the image over which the light is spread. But the linear dimensions of the image are proportional to f, the focal length of the lens, and so the *area* of the image is proportional to f^2. Hence, other things being equal, it is the ratio $\dfrac{A}{f^2}$ which determines the time of exposure.

Figure 546 represents a symmetrical or "rectilinear" lens consisting of two similar achromatic lenses, symmetrically placed, and having the diaphragm half-way between them. It will be observed that in this case the oblique pencil passes as much below the center of the front lens as above the center of the back lens, so that the beam is as much bent by one as by the other and emerges parallel to the incident pencil. This

tends to cause straight lines in the object to be reproduced as straight lines in the image.

909. Distortion of Images. The image of a grating with equal square openings may be distorted in either of the two modes shown in figure 547. The first or barrel-shaped distortion is seen when the center of the image is magnified relatively more than the outer portions, while the other form
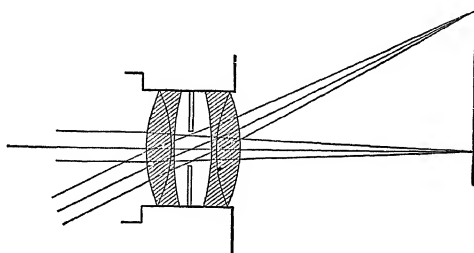
FIG. 546

of distortion is caused by the greater relative magnification of the parts away from the center.

Either of these modes of distortion may be produced in projecting the image of the grating with the same lens, the form of distortion being determined by the mode of illumination. For let G (Fig. 548) be the grating and L the lens, then if the grating is illuminated by a beam of nearly parallel
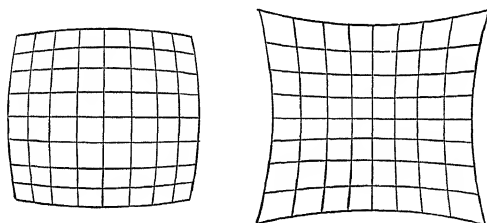
FIG. 547

light, as direct sunlight, the light from the upper part of the grating will pass through the upper edge of the lens L, and being bent down too strongly in consequence of the spherical aberration of the lens (§ 895) will come to focus at P farther from the center than P' and will thus cause the distortion shown in the second diagram of figure 547. If, on the other hand, by means of a convergent lens the illuminating beam of light is converged so strongly that rays from the top of G are refracted by the bottom of the lens L, as shown in figure 549, the focus P will be too near the center and the distortion will be barrel-shaped.

It is clear that *the least spherical aberration and distortion will be secured when the illuminating lens converges the light toward the center of the lens L,* as shown in the diagram of the magic lantern, figure 550.

It is this same kind of spherical aberration which causes barrel-shaped distortion in the photographic image when the diaphragm is placed (outside)
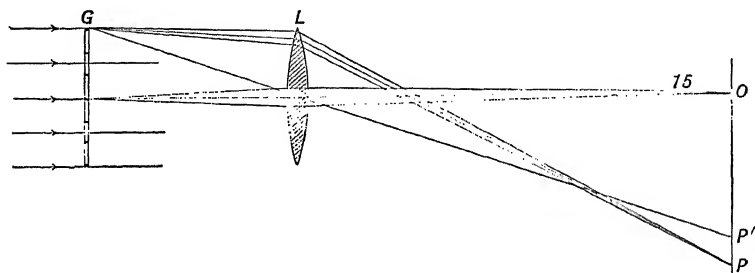
FIG. 548.   Parallel illumination.   Pin-cushion-shaped distortion

in front of the lens; while if the diaphragm is behind the lens the opposite form of distortion results.

910. Projecting Lantern. The optical system of the *magic lantern*, *stereopticon*, or *projecting lantern* is shown in figure 550.
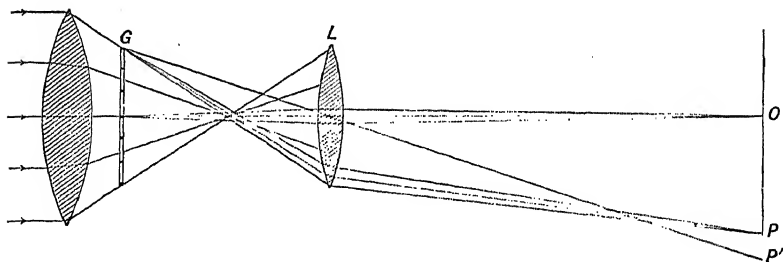
FIG. 549.   Barrel-shaped distortion

It consists simply of a front lens or objective L which forms a real image of the slide S on the screen at S'; and an illuminating system which consists of the source of light at E and the condensing lens C which converges the light through the slide S toward the center of the lens L.

Since the screen S' is usually at a considerable distance, the distance from the slide S to the lens L is nearly the focal length

of the lens or lens combination. So that the width of the image on the screen is to the width of the slide as the distance of the screen is to the focal length of the front lens L. Hence when the lantern is to be at a great distance from the screen a long-focus front lens should be used to prevent the image from being too large and dim.

The front lens is usually a combination of two lenses to secure flatness of field and freedom from color and distortion. The condensing lens consists of two plano-convex lenses with their
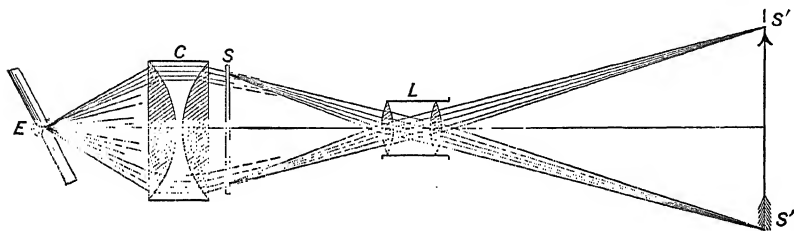
FIG. 550. Projecting lantern

convex surfaces almost touching. If the upper portions of the two condensing lenses are thought of as prisms, it will be noticed that with this construction each is nearly in the position of *minimum deviation* for the pencil of light passing through it; for the incident and emergent pencils make somewhat nearly equal angles with the two surfaces of each of the two lenses. Such an arrangement makes the spherical and chromatic aberration very much less than if the lenses had been placed with their flat faces together, which would make practically a single double-convex lens.

911. Projecting Microscope. For the projecting of microscopic objects the lens L must have a very short focal length to secure the requisite magnification. The object must also be intensely illuminated and so a second short focus condensing lens C' is introduced which converges the light from E to a bright focus at the slide S which is to be illuminated. To diminish the heat of the focus at S, which might ruin the slide, a tank of water 3 or 4 in. thick is introduced between C and C'. Water absorbs strongly the very energetic radiations whose wave lengths are too long to affect the eye, though the visible radiation or

light is not sensibly weakened. An ordinary microscope may be used in this way for projection either with or without the eye-piece by turning the instrument into the horizontal position
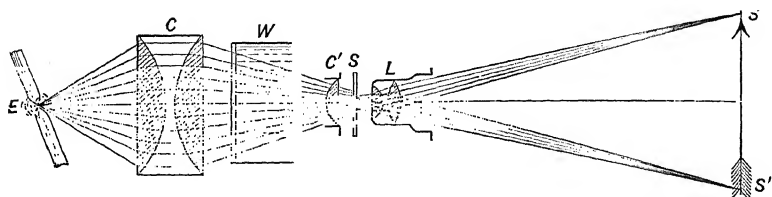
FIG. 551

and converging a beam of sunlight on the slide *S* by a convergent lens of 8 or 10 in. focal length.

912. The Eye. The human eye is nearly spherical in shape, having an outer wall of firm textured substance of which the transparent front is known as the *cornea*. Immediately back of the cornea is the *iris*, a variously colored membrane, having a round opening in its center called the *pupil*. The pupil contracts in bright light and dilates in the dark, the iris acting as a diaphragm to regulate the light admitted to the eye. Back of the pupil is the crystalline lens, of rather dense transparent tissue formed in layers and densest at the center. The space between
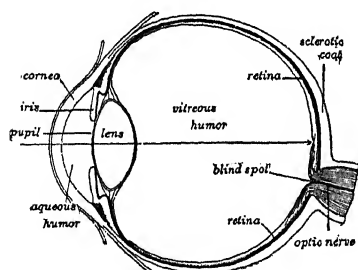
FIG. 552

the crystalline lens and the cornea is filled with a clear watery substance, the *aqueous humor*, while the main interior cavity back of the crystalline lens is filled with a transparent jelly-like substance, the *vitreous humor*. At the back of the eye, forming the inner coating of the outer wall, is the *retina*, a highly organized black membrane the surface of which is covered with minute structures called *rods and cones* in which the fibers of the optic nerve terminate.

Rays from external objects are focused on the retina by the action of the crystalline lens and other refracting portions of the eye. An image is formed on the retina just as the image in a

photographic camera is formed on the plate, and each portion of the retina thus receives a stimulus exactly corresponding to the illumination of the particular part of an external object which has its image at that point, and this stimulus of the optic nerve causes the corresponding sensation of brightness and color.

Of course the image formed on the retina is inverted, but we do not see the image inverted; there is simply a correspondence between the retina and external directions, such that when light falls on a spot on the retina it excites a sensation which we describe by saying that it is bright in the corresponding direction.

At the center of the retina and just opposite the pupil and crystalline lens is a spot where the retina is much more highly developed than elsewhere, and to see objects distinctly their images must be formed on that spot. If the eye is directed at a particular point on a printed page only the words close to that point are seen distinctly.

Where the optic nerve enters there is a *blind spot* in the retina. To verify this, make a small black spot on a sheet of white paper and covering the left eye look with the right eye at a point about one-fourth as far to the left of the spot as the latter is distant from the eye and the spot will disappear.

913. Accommodation. A normal eye can change its focus or *accommodate* so as to form on the retina a distinct image either of distant objects or of those as near as about 8 in. This is brought about by muscles attached to the crystalline lens by which it is made flatter for distant objects and more convex for nearer ones.

914. Short and Far Sight. When the lens of the eye is too convex, objects at ordinary distances are focused *in front of the retina*, so that the image on the retina itself is out of focus and blurred. In such a case objects can be seen distinctly only if held very near the eye, and the person is said to be *short*-sighted, or *myopic*.

If, on the other hand, the lens of the eye is too flat the image of a near object will be formed *back of the retina*, so that indistinct vision results. In such a case it may be that only distant objects can be seen distinctly or it may not be possible to see distinctly at any distance, and the person is said to be *presbyopic*, or *far*-sighted.

915. Spectacles. If the lens of the eye is too convex as in case of short sight, a divergent lens may be used to correct the defect, while if the lens of the eye is too flat, as in far sight, a convergent lens must be used.

916. Astigmatism. An eye is said to be *astigmatic* when a point of light, as a star, is seen as a short bright line, the direction of which is called the axis of astigmatism. In case of astigmatism all the lines in such a diagram as figure 553 will not appear equally distinct, but those in the direction of the axis of astigmatism will be sharply defined while those at right angles to them will appear broadened and blurred. This defect is caused by the lens of the eye having ellipsoidal instead of spherical surfaces and is corrected by the use of cylindrical lenses.
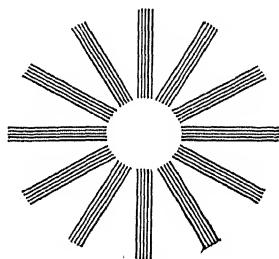
FIG. 553

917. Distance of Distinct Vision. The nearer an object is brought to the eye the larger will be the dimensions of its image on the retina and the more detail will be brought out, provided it is not brought so near that the eye cannot properly focus the image.

A distance of about 10 in. or 25 cms. is the *normal distance of most distinct vision*.

918. Magnifying Glass. A convergent lens produces a virtual and enlarged image of any object placed slightly nearer to the
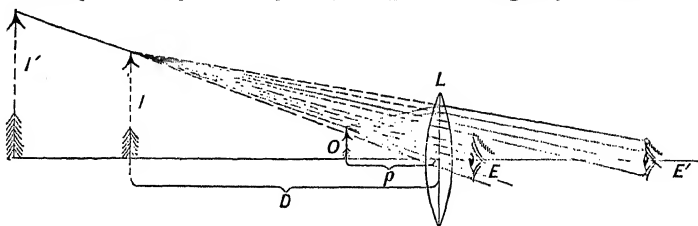
FIG. 554

lens than its principal focus. In the diagram it will be seen that light from each point of the object *O* after passing through the lens *L* comes to the eye as if it had come from the image *I*. The object is therefore seen enlarged or magnified.

In using a magnifying glass *the eye should be placed close to the lens as at E and the object then be brought up until its image is seen distinctly*, for in that case the rays from all parts of the object come to the eye through the *central part of the lens*, and are subject to the least spherical and chromatic aberration; whereas if the eye were placed at E' rays coming to the eye from the point of the arrow would be refracted near the edge of the lens and consequently there would be distortion of the image as well as other aberrations.

The angular dimension of the image as seen by the eye placed at E close to the lens is practically the same whether the image is formed at I or I' or at a still greater distance. Therefore the *apparent size* of the image and consequently the magnification is substantially unchanged.

The magnifying power of the lens is the ratio of the length of the image I, formed at the distance of distinct vision D, to the length of the object O; for in order to see the object distinctly without the lens it must be placed at the distance D from the eye.

To determine the magnifying power, let f represent the focal length of the lens and let p and D represent the distances from the lens of O and I, respectively, then by definition the magnifying power of the lens $= \dfrac{I}{O}$

but

$$\frac{I}{O} = \frac{D}{p}$$

and from the general lens formula (§ 888) we have

$$\frac{1}{p} - \frac{1}{D} = \frac{1}{f}.$$

Multiplying through by D this becomes

$$\frac{D}{p} = 1 + \frac{D}{f},$$

therefore the magnifying power of a lens $= 1 + \dfrac{D}{f}$ where D is the distance of most distinct vision and f is the focal length of the lens.

919. Compound Microscope. For the highest magnification the compound microscope is used, the optical system of which is shown in figure 555. The object O to be magnified is placed just outside the focus of the *short-focus lens A*, called *the objective*, which forms a real image at I. Back of this image is placed a magnifying glass at a distance slightly less than its focal length so that it forms a virtual image of I at I' which may be seen by the eye at E.

Since O is nearly at the principal focus of the lens A, the image I will be as many times greater than the object O as the
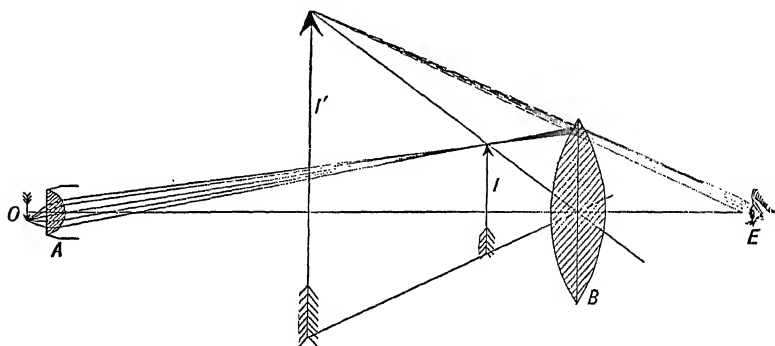
FIG. 555.   Compound microscope

distance AI is greater than the focal length of A. The distance AI in an ordinary microscope is about 150 mm., so that if the focal length of the objective is 5 mm. the image I will be 30 times as large as the object, and if the eye-piece or lens B has a magnifying power 10, the power of the combination is $30 \times 10 = 300$ diameters.

Many microscopes have a "draw tube" by which the distance between the objective and eye-piece may be increased. The effect of this is to increase the magnifying power of the instrument in the same ratio.

920. Eye-pieces and Micrometer. The *eye-piece* or *ocular* of a microscope usually consists of a combination of two lenses instead of the simple lens B shown in the previous diagram. Two forms are in use. One of these, the *Huygens* or *negative eye-piece*, consists of two convergent lenses, a *field lens F* and *eye lens E*, whose focal lengths are in the ratio 3 to 1, and the distance between them is twice the shorter focal length. It is equivalent

to a single lens whose focal length is $1\frac{1}{2}$ times that of the eye lens, but it has the advantage of being partially achromatic if both lenses are made of the same kind of glass, and gives less distortion and aberration than a single lens.

In this eye-piece the rays from the objective must fall on the field lens *before* coming to a focus at *I*, as shown in the figure.

For rough measurements this eye-piece may be provided with a scale
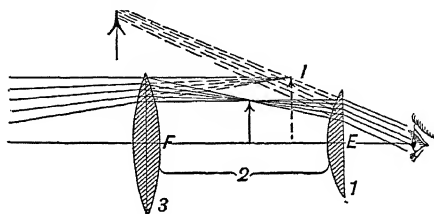
FIG. 556. Huygens' or negative eye-piece. Focal length = $1\frac{1}{2}$

ruled on glass, known as an *eye-piece micrometer*, which is fixed between the two lenses so that it coincides with the image formed by *F*.

But for more exact measurements the microscope is provided with a *micrometer* (in which a cross-hair is moved by a screw having a graduated head) placed at *I* where the image is formed by the objective, and the eye lens must be placed *back* of this point so as to magnify image and cross-hairs together.

For this purpose the *Ramsden* or *positive* eye-piece is used. This form of ocular consists of two convergent lenses, a field lens and an eye lens
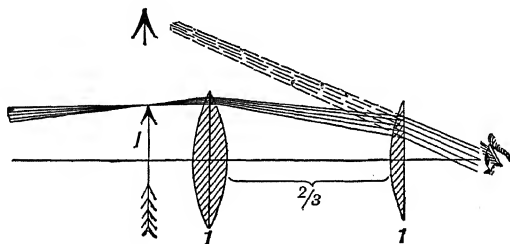
FIG. 557. Ramsden's or positive eye-piece. Focal length = $\frac{3}{4}$

of equal focal length, the distance between them being two-thirds of the focal length of either lens. This eye-piece is placed back of the image *I* just as though a single lens were used. It is equivalent to a single lens whose focal length is $\frac{3}{4}$ that of its component lenses, and is called *positive* because it can be used, like an ordinary simple magnifying glass, to magnify any small object. It is nearly achromatic and may give a flat field with little aberration.

921. Microscope Objectives. The object-glass of a microscope usually consists of a nearly hemispherical front lens of crown glass with its flat face outward, having one or more achromatic combinations mounted back of it as shown in the figure. The curvatures of the lenses and their distances apart are calculated so as to give an image as free from aberration as possible. High powers are corrected for a particular tube length and thickness of cover-glass, and to obtain the best results these conditions must be satisfied. If between the front lens of the objective and the cover-glass a drop of oil is introduced having the same index of refraction as the glass, it is as though the object were imbedded in the glass of the front lens. In this way loss of light by reflection from the glass surfaces is avoided, no correction for the thickness of the cover-glass is required, and the brightness and definiteness of the final image are increased. But to secure these results the objective must be especially designed for this use. Such objectives are known as *oil* or *homogeneous immersion* lenses, and may have as short a focal length as $\frac{1}{12}$ or $\frac{1}{16}$ in.
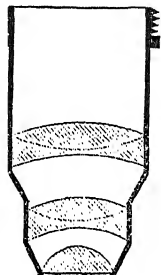
Fig. 558

922. Numerical Aperture. The size of the pencil of light transmitted through the microscope from a single point of the object has an important influence on the defining power (§ 969). *The ratio of the radius of the largest cross section of such a pencil to the focal length of the objective is known as its numerical aperture,* and, other things being equal, the resolving power of an objective is proportional to its numerical aperture.

923. Telescopes. In telescopes a convergent lens, known as the *object-glass,* or a concave mirror, is used to form a real image of a distant object, and this image is then magnified by a suitable eye-piece. *Since the object is distant the image is formed at the principal focus of the object-glass or mirror, and consequently to have a large image and great magnifying power the object-glass must have long focal length.* There are three kinds of refracting telescopes, Galileo's, the astronomical, and the terrestrial forms.

924. Galileo's Telescope. In Galileo's telescope a concave or divergent lens is used as the eye-piece. This lens is placed so that rays from the object-glass meet it *before* forming the image I as shown in the figure. If the distance from L to I is slightly greater than the focal length f of the eye lens, rays approaching the point of the image at I will be bent upward and made to diverge as if from I'. An enlarged virtual image is thus formed which may be seen if the eye is placed so as to receive the emer-

gent pencil. It will be observed, however, that the pencils of light from all points in the object come to the eye as if intersecting at S, so that it is as though the virtual image were seen through a small opening at S which restricts the field of view, only so much being seen in any one position of the eye as is in line with the eye and S. *The smallness of the field of view is the great defect in this form of telescope and causes its use to be restricted to low-power instruments, such as the ordinary opera-glass.* Its advan-
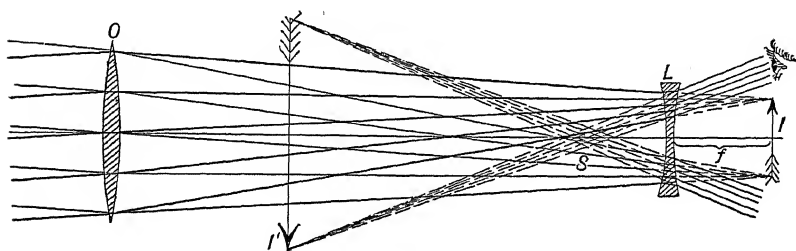
FIG. 559.   Galileo's telescope

tages are that it is shorter than the other forms of telescope and gives an erect image of the distant object.

925. The Astronomical Telescope. In this instrument the eye-piece is a convergent lens, or an equivalent Huygens or Ramsden eye-piece as in the microscope (§ 920). The eye lens, as shown in the diagram, is placed nearly its own focal length back of the image at I formed by the object-glass, so that a virtual enlarged image of I is formed at I'. It will be noted that the pencil of rays from the lower part of the distant object comes to the eye as if from the upper part of the image at I'. In this instrument, then, the image is inverted, and it is therefore used chiefly for astronomical observation. The various pencils of rays coming to the eye from different points in the virtual image all intersect at S, forming a bright spot known as the *eye spot.* If the pupil of the eye is held at this point all parts of the virtual image can be seen simultaneously, and the field of view is large, being limited only by the size of the lens L.

926. Magnifying Power of a Telescope. The semi-diameter of the distant object as it would be seen without the telescope subtends the angle COB or IOD (Fig. 560) where CO is a ray

coming from the middle of the object and BO is a ray from its
edge; while the angle subtended by the corresponding part of the
image seen in the telescope is $I'SD$ or ILD.   The magnifying
power of the instrument is therefore the ratio

$$\frac{tan\ ILD}{tan\ IOD}.$$

Since OD is equal to F, the focal length of the object-glass, and
DL is nearly equal to f, the focal length of the eye lens, we have

$$tan\ IOD = \frac{ID}{F}, \qquad \text{and} \qquad tan\ ILD = \frac{ID}{f},$$

$$\text{therefore} \qquad \frac{tan\ ILD}{tan\ IOD} = \frac{F}{f}$$

or the magnifying power of an astronomical telescope is equal to
the number of times that the focal length of the eye-piece is con-
tained in the focal length of the object-glass.
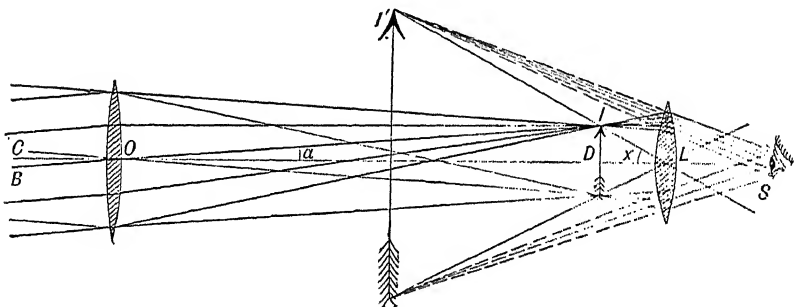
FIG. 560.   Astronomical telescope

It may also be shown that the magnifying power is the ratio of
the diameter of the object-glass to the diameter of the eye-spot S,
for the latter is the image of the object-glass formed by the lens L.

It will be shown later (§ 969) that the *defining power* of a
telescope is proportional to the diameter of the object-glass,
supposing it to be perfect.

Hence for detecting close double stars or for investigating
minute details on the surface of sun or planet where high powers
must be used, it is necessary to employ a telescope of large

aperture. Large lenses are used also on account of their light-gathering power in observing faint objects, such as nebulæ.

927. Terrestrial Telescope. To obtain a large field of view and at the same time an erect image of distant objects the terrestrial telescope is used.

This instrument is like the astronomical telescope except that there are two additional convergent lenses, E and E', introduced
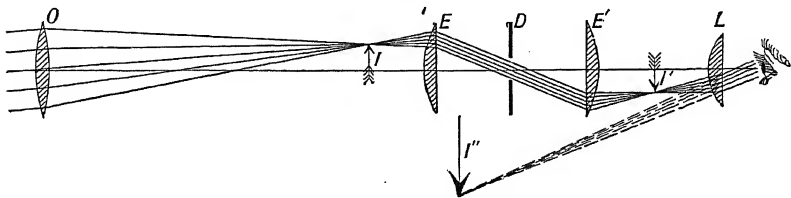
FIG. 561. Terrestrial telescope

between the object-glass O and the eye lens L as shown in the diagram. These lenses invert the image I, forming another real image at I' with the point of the arrow downward as in the distant object. This image is magnified by the eye-piece at L, which forms the enlarged erect virtual image I''. In the ordinary spy-glass the lens L is replaced by a Huygens eye-piece, making four lenses in all besides the object-glass.

At D a diaphragm is introduced having a hole in the center just large enough to transmit the pencils of light which intersect at
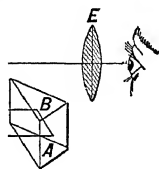
FIG. 562. Path of rays in prism binocular

that point. This serves to stop any stray light which may be reflected from the sides of the tube.

928. Prism Binoculars. Field glasses which combine high magnifying power and large field of view are now made according to the plan shown in figure 562.

The beam of light from the object-glass enters a right-angled

glass prism and after two internal reflections, at A and B, is completely reversed in direction and travels back to a second prism, placed at right angles to the first, where it is again totally reflected, at C and D, and turned back again toward the eye lens E.

The reflections in the two prisms secure an erect image without using the reversing lenses of the ordinary terrestrial telescope; for one prism interchanges the two sides of the image, while the other makes it upright, thus restoring it completely to its natural position.

Also, on account of the length of the path of the rays from the object-glass to the eye lens, the focal length of the object-glass may be three times as great as in the ordinary field glass, and the magnifying power correspondingly increased.

929. Reflecting Telescopes. Instead of the object-glass of a telescope, a long-focus concave mirror may be used. The arrangement shown in figure 563 was used in Sir William
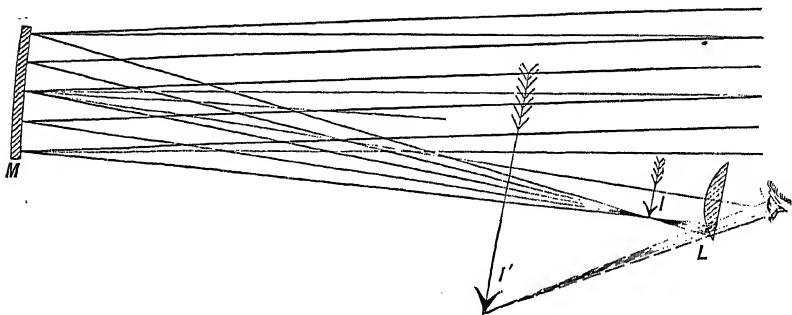
FIG. 563. Herschel's form of reflecting telescope

Herschel's great telescope, the mirror M being slightly inclined so that the eye-piece and the observer's head were not in the line of the rays falling in the mirror. In small reflecting telescopes the rays converging toward the image I may be reflected out sideways to the eye-piece by a small mirror placed directly in front of the large mirror, as was done by Newton. To obtain a perfect image free from spherical aberration the mirror must be parabolic instead of spherical. A mirror has the advantage of forming an image *free from chromatic aberration* since all wave lengths of light are reflected alike.

A telescope mirror is called a *speculum,* and may be made of an alloy called speculum metal which takes a fine polish and does not readily tarnish. Specula are now, however, usually made of glass, as this is harder and less dense than speculum metal. The surface is ground and polished to the required shape and then silvered.

PROBLEMS

1. What is the magnifying power of a simple convergent lens of focal length 5 cms.? Take 25 cms. as distance of most distinct vision.

2. What is the magnifying power of an astronomical telescope in which the focal length of the object-glass is 12 ft., while that of the eye-piece is 1.5 in.?

3. A compound microscope has an objective of 1 in. focus, the first image is formed 5 in. back of the objective and is magnified by an ocular of 2 in. equivalent focus. Find the magnifying power of the combination.

4. When a telescope is pointed at the sun, how should the eye-piece be placed to give a real image of the sun on a screen fixed back of the telescope? In this case is the image formed by an astronomical telescope erect or inverted?

5. In a projection apparatus it is desired that the pictures shall be 10 ft. wide on a screen 40 ft. distant, when the slides are 3 in. wide. What must be the focal length of the projecting lens used?

6. By means of a microscope objective, a scale having 50 lines to the millimeter is projected upon a screen 9 meters distant, and the distance between the lines in the image on the screen is 4 cms. What is the focal length of the objective used?

7. A certain binocular field glass has a power 8 and diameter of objective 25 mm. while another of power 6 has an objective 21 mm. in diameter. Which will give the brighter image and in what ratio?

REFERENCE

J. T. Taylor: *Optics of Photography.*

Analysis of Light

The Spectrum

930. The Spectrum. It has been seen that when light passes through a prism it is spread out in a colored band shading from red to violet and called the spectrum, showing that a beam of white light is complex and made up of different kinds of light

which are separated by the prism in consequence of their differ-
ent refrangibilities. These lights also differ in the color sensa-
tions which they excite, the least refrangible being the red rays
while the most refrangible are the violet.

It will be shown later that the physical property which deter-
mines the refrangibility of a ray of light is *wave length;* so that
*in forming a spectrum we are really spreading out the light in the
order of wave lengths, the longest waves being at the red end of the
spectrum and the shortest at the violet end.*

931. Pure Spectrum. To obtain a complete analysis of light
there must be no overlapping of different kinds, but each must
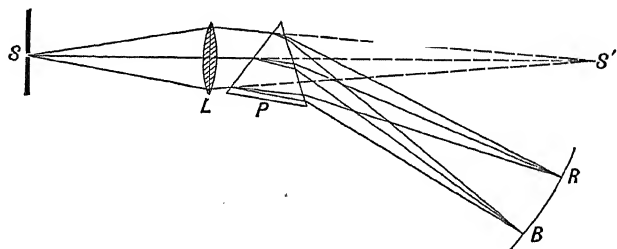
FIG. 564

be separated by the prism from every other. To accomplish
this the following arrangement may be employed.

The light to be examined enters through a narrow slit *S*,
which is parallel to the edge of the prism and therefore perpen-
dicular to the plane of the paper in the diagram. A converging
lens *L* is so placed that it would bring the light to focus and
form at *S'* an image of the slit if the prism were not interposed.
By the action of the prism, however, the light is refracted down-
ward and if there were only one kind of light present the whole
beam would be equally refracted and the bright image of the
slit would be formed at *R*, say, instead of at *S'*, but without any
change in color. If, however, there were in the original beam
two kinds of light which were differently refracted, there
would then be formed two images of the slit, one at *R* and one at
B. And since light waves that are refracted differently also act
differently upon the eye, the images at *R* and *B* will be of differ-
ent colors. But if the original beam contained waves of every

conceivable degree of refrangibility within certain limits, there would be an infinite number of images of the slit with no separation between them and even overlapping, producing a continuous band of color, shading from one extreme to the other. When the slit is narrow so that the amount of overlapping is small the spectrum is said to be *pure*.

932. Fraunhofer Lines. When the *sun spectrum* is formed as above described, using a narrow slit so as to produce a pure spectrum, a large number of dark lines are observed which cross the spectrum parallel to the slit, showing that sunlight does not
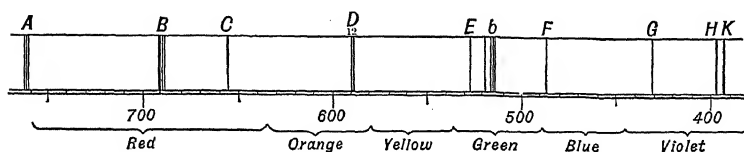
FIG. 565. Fraunhofer lines. The wave lengths are given in millionths of a millimeter

contain all kinds of light, but that certain wave lengths are lacking, and consequently no bright images of the slit are formed at the corresponding points of the spectrum. These dark lines characteristic of sunlight were first carefully studied by Fraunhofer and are known as the *Fraunhofer lines*. Some of the most prominent of them are designated by letters of the alphabet, and furnish convenient points of reference in the spectrum, the *A* line being almost at the limit of visibility in the red while the *H* line is near the extreme violet.

933. Analysis of Light by Spectroscope. For the more exact analysis of the light from any source a *spectroscope* is used. The main features of this instrument are indicated in figure 566. Light from the source to be studied enters the narrow slit *S* at the focus of the collimating lens, diverges through that lens and then passes as a parallel beam to the prism *P* where it is refracted and dispersed. After passing the prism the beam enters the telescope and forms a sharply defined spectrum at the principal focus of the object-glass of the telescope, a separate image of the slit being formed by each wave length of light present. This spectrum is magnified by the eye-piece of the telescope.

An illuminated scale is also sometimes mounted so that light from the scale, reflected at the second face of the prism, enters the telescope and forms an image of the scale along with the
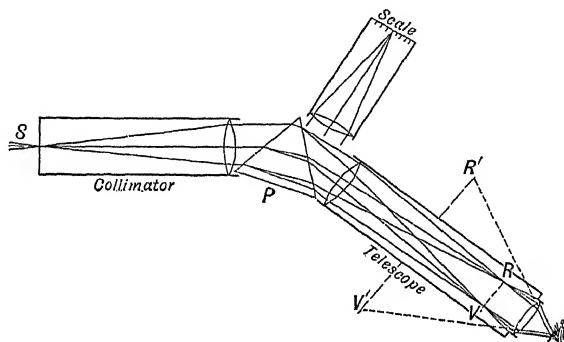
Fig. 566.　Spectroscope

spectrum at RV. The observer can by this means locate a definite line in the spectrum by its position on the scale.

To compare the spectra from two separate sources a *comparison prism* is sometimes used. This is a small total reflecting prism covering one-half of the length of the slit. Light from one source enters the spectroscope directly through the uncovered half of the slit, while the other source is so placed that its light is reflected by the prism through the other half of the slit. The two spectra are formed side by side as shown in figure 567,
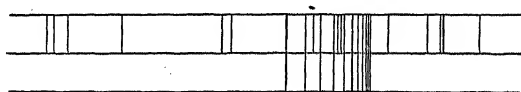
Fig. 567.　Diagram of a carbon band below; the upper spectrum has also lines due to other substances

the wave lengths of lines that match in the two spectra being the same.

To secure greater dispersion than can be obtained with a single prism a train of several prisms may be used, so placed that the light passes through them in succession before entering the telescope.

934. The Complete Spectrum. Instead of observing the spectrum by the eye, it may be received on a photographic plate and photographed. When this is done it is found that there are rays beyond the violet light of the visible spectrum and still shorter in wave length, which act on the photographic plate but are invisible to the eye. These rays are called *ultra violet* light.

Another mode of examining the spectrum is by means of a delicate thermopile or bolometer by which the heating effect of the rays at each point in the spectrum may be determined. Professor Langley perfected an apparatus of this sort in which
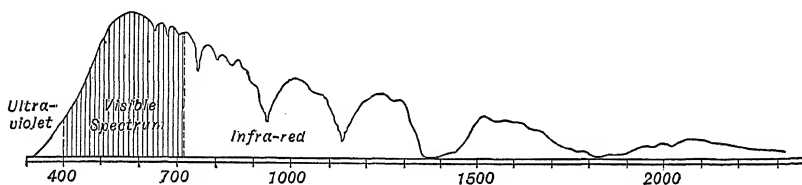
Fig. 568. Energy spectrum, after Langley. Curve showing distribution of energy in the sun spectrum, plotted by wave lengths in millionths of a millimeter

the bolometer filament was carried along by clockwork from one end of the spectrum to the other, while at the same time a beam of light reflected from the mirror of the galvanometer fell on a moving strip of sensitized paper recording the deflection of the galvanometer for every point in the spectrum. A curve obtained in this way showing the heating effect of different parts of the sun spectrum is shown in figure 568.

Such a study shows that beyond the red end of the sun spectrum there is an invisible region where the radiation has great energy or heating power. This region is known as the *infra red*, and its waves are longer than those of the visible red. The shaded area in the diagram represents the distribution of energy in the *visible* spectrum.

If a plate of clear glass is held between the face and an open fire the warmth of the radiation is greatly cut off, though the visible radiation is almost completely transmitted. The change is due to the power of glass to absorb the infrared radiation. When the spectrum of fire light is taken by a bolometer first

directly and then through a pane of glass a comparison of the two curves shows just what wave lengths are absorbed by the glass.

There is no reason to suppose that the infra red or ultra violet light is different except in wave length from ordinary visible light. The luminous body is to be thought of as giving out waves of different lengths from the extreme infra red to the extreme ultra violet; all of these waves have *energy* and consequently have heating effect, but only certain wave lengths can affect the eye. The luminous effect of the part called the *visible spectrum* is due to a peculiar response which waves of certain frequencies excite in the eye and which we call the sensation of light. The photographic effect of certain waves also depends on the responsiveness of the substances acted on, to waves of a special frequency of vibration.

The earlier writers speak of heat rays, luminous rays and chemical or actinic rays as though there were three different kinds of radiation. In the use of these terms care should be taken to guard against such a misconception.

935. Absorption in Spectroscopes. A glass prism absorbs strongly long radiation beyond the visible spectrum and also the shorter waves in the ultra violet. Hence *to study the energy spectrum in the infra red* a prism of rock salt, fluorspar, or sylvite must be used.

Of these sylvite transmits waves up to 0.025 mm. in length, while the range of rock salt and fluorite is somewhat less, 0.020 mm. being the limit with the former and 0.011 mm. with the latter. For very long waves, more than 0.050 mm. in length, quartz is transparent, though it absorbs strongly the shorter wave lengths transmitted by rock salt and sylvite.

In the study of *the short waves in the ultra violet* part of the spectrum quartz prisms and lenses are used. But the shortest waves are so strongly absorbed by air that they can be detected and studied only when the apparatus is in a vacuum. In this way Lyman and Schroeder have photographed light waves shorter than 0.0001 mm.

936. Kinds of Spectra. There are three different kinds of spectra: *continuous, bright-line,* and *absorption spectra.*

A continuous spectrum contains all wave lengths between

certain limits, and if visible appears to the eye as a continuous band of color shading from one end to the other. Hot solids and liquids give rise to continuous spectra.

Bright-line spectra are obtained when the source of light gives out only certain definite wave lengths. Each wave length gives a bright line in the spectrum, the intensity of which depends on the energy of the corresponding mode of vibration. Gases and vapors when rendered incandescent by heat or by the electric discharge give bright-line spectra.

These spectra are highly characteristic, every known substance having a different spectrum which depends to some extent on the method used to make it luminous. The investigation of substances having lines in their spectra that could not be attributed to any known element has led to the discovery of a number of new elements. The identification of substances by their spectra is known as *spectrum analysis*.

Absorption spectra are obtained when the light from some source which would give a continuous spectrum is made to pass through some absorbing medium and then analyzed by the spectroscope.

Spectra due to absorption by solids or liquids usually show broad absorption bands shading off at both edges, while absorption by gases and vapors gives rise to sharply defined black lines in the spectrum, showing that only certain special wave lengths are absorbed. The Fraunhofer lines in the sun spectrum are produced in this way.

937. Production of Bright-line Spectra. If a loop of platinum wire is dipped into a solution of some salt of sodium, potassium, barium, lithium, or strontium, and is then held in a very hot non-luminous flame, like that of a Bunsen burner, the flame becomes colored, bright yellow in case of sodium, red by lithium and strontium, and violet by potassium; and when the colored light is examined by the spectroscope, the spectrum is found to consist of certain characteristic bright lines. Usually certain of these lines are particularly bright and prominent and determine the characteristic color, but a higher temperature will often bring out others of less intensity.

To obtain the spectra of substances such as iron, copper, zinc, etc., which volatilize only at very high temperatures, the electric arc may be used, the substance to be studied being introduced

into a cavity in the end of one of the carbons. In this case the
spectrum of the carbon arc is also present and its lines must be
distinguished from those due to the substance which is intro-
duced. A better method is to obtain the arc between terminals of
the pure metal which is to be studied, though there are
practical difficulties in carrying this out.

Still another method of obtaining spectra of these
substances is by the *spark discharge* between terminals
made of the substance whose spectrum is to be de-
termined. The sparks may be produced by an induction
coil and intensified by the use of a Leyden jar having
its inner coating connected to one terminal and its outer
coating to the other. If the discharge is sufficiently
intense, lines are observed which are characteristic of
the substance of which the terminals are made.

The spectrum obtained by volatilizing a substance
in the *electric arc* is generally somewhat different from
its spark spectrum.

To study the spectrum of a gas it may be enclosed
in a tube such as shown in figure 569, known as a
Plücker tube, and made to glow by sending the electric
discharge from an induction coil between its two aluminum
electrodes. The two bulbs at the ends containing the electrodes
are connected by a capillary tube in which the discharge is con-
centrated and intensely luminous. It is this capillary portion
which is placed in front of the slit of the spectroscope.

FIG. 569
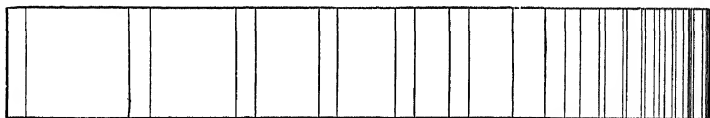
FIG. 570. *B* group due to oxygen in the earth's atmosphere.
Sun spectrum

938. Fluted Spectra. In some cases the spectrum presents
the appearance of a series of shaded bands or *flutings*, well
shown in the spectrum of nitrogen. But when examined with
a spectroscope of high dispersive power these flutings are each
seen to be made up of a regular series of bright lines crowded
closely together near the bright edge of the fluting and at dis-

tances apart which increase regularly from the bright to the faint edge of the fluting, somewhat as shown in figure 570 which represents a remarkable group of lines in the red end of the sun's spectrum, the B group due to oxygen in the earth's atmosphere.

939. Source of Light Waves. There is little doubt that the source of nearly all visible and ultraviolet violet light is within the atom itself although some of the long waves of the infrared are of molecular origin. It was for many years believed by physicists that each vibration in the light spectrum was produced by a corresponding vibration of an electron within the atom. It can be shown according to Maxwell's theory (§ 1006), that a vibrating electric charge should produce a train of electromagnetic waves of frequency equal to the frequency of vibration of the charge which produces them. The combined work of experimental and mathematical physicists in the last ten years and more has shown that radiation from atoms is not produced by vibrating electrons but that light vibrations arise from certain energy changes within the atom and which differ for every different kind of atom. These energy shifts are a quantum phenomenon and have already been touched upon in § 801. The nature of radiation is more fully explained in the final chapter of this book entitled "Radiation and the Quantum Theory." The physical laws of the atomic processes involved in radiation are very different from those of classical physics to which we are accustomed in every-day life. The study and the explanation of the nature of radiation from atoms occupy many of the ablest physicists of the present day.

The atomic energy shifts which produce light waves are chiefly excited either by molecular agitation of high temperature or by electrical excitation such as that produced in a Geissler tube or in an electrical spark, or by excitation from radiation.

The different energy changes which produce the different spectrum lines are comparatively few for gases where the molecules are comparatively far apart, so that spectra consisting of separate lines are produced by them. For liquids and solids the molecules and atoms interfere with one another, which interference results in such an enormous number of possible energy changes that a continuous spectrum is produced.

Single atoms by themselves produce only certain lines each

time an energy change takes place. The complete spectrum of a radiating substance at a given temperature arises from enormous numbers of atoms, each emitting some one of the characteristic frequencies simultaneously, so that the total effect is the complete spectrum of that substance for that temperature.

940. Explanation of Fraunhofer Lines. When the sun's spectrum and that of some elementary substance are photographed beside each other on the same plate it is found in many cases that the bright lines in the spectrum of the substance exactly match,
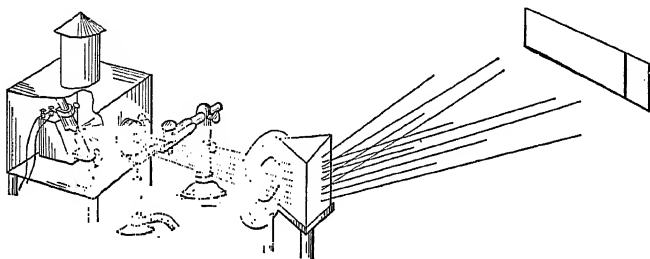
FIG. 571. Absorption by sodium vapor

line for line, certain of the dark Fraunhofer lines in the sun's spectrum. For example, the two yellow sodium lines exactly coincide with the two D lines in the solar spectrum.

The explanation of these dark lines in the solar spectrum was given by the German physicist, Kirchhoff, who showed that they might be caused by the absorption of light coming from the deeper layers of the sun in passing through the cooler vapors in the sun's outer atmosphere, and announced the principle that *a vapor or gas will absorb most powerfully light of the same wave lengths as the light which the same gas or vapor gives out when it is itself the source of radiation.* This principle is illustrated by the following experiment. Form a pure continuous spectrum as described in § 931, using as the source of light the glowing positive carbon of the electric arc, and volatilize a fragment of metallic sodium just below and close in front of the slit by means of an alcohol or Bunsen flame, or, better, put a fragment of metallic sodium in a little cavity in the lower carbon of the arc itself. The light from the arc passes through the dense cloud of sodium vapor and a dark line appears in the orange-

yellow of the spectrum, just where bright lines are found in the spectrum of sodium, showing that the waves absorbed are of the same wave length as those given out by glowing sodium vapor.

The black line in this case is not strictly black as it is illuminated by the radiation from the sodium vapor, but this is so much less intense than the direct radiation from the arc that it appears black by contrast. It is clear, therefore, that to produce black absorption lines *the absorbing vapor must be colder than the luminous source, or at least its direct radiation must be less intense than that which it absorbs.*

Of course, in *a steady state of things a mass of vapor in the atmosphere of the sun must be radiating just as much energy as it absorbs,* otherwise it would be growing hotter or colder; but the radiation which it absorbs comes to it mainly in one direction, while it radiates equally in every direction; therefore the radiation which it sends to the earth must be much less intense than that which it intercepts.

941. Astronomical Spectroscopy. To examine the spectrum of a star or of a particular portion of the sun it is only necessary to form on the slit of a spectroscope an image of the object to be examined. For this purpose the eye-piece of the telescope is removed and a spectroscope is mounted so that its collimator is in line with the axis of the telescope and its slit at the principal focus of the object-glass.

In stellar spectroscopy no slit is required since the image of a star is a mere point of light, and the spectra of neighboring stars may be simultaneously photographed on the same plate. But it is also a consequence of the smallness of the stellar image that the spectrum of a star is a mere line of light too narrow to show well the spectrum lines. The breadth of the spectra may be increased by using a cylindrical lens; or the motion of the telescope may be so regulated that the spectra shift slowly on the photographic plate, at right angles to their lengths, so that each leaves a broad trace on the plate.

942. Doppler's Principle in Spectroscopy. Doppler's principle, by which the apparent pitch of a sounding body is raised when it is moving toward the ear and lowered when it is receding (§ 315), has also an important application in case of light waves. While a luminous body is moving *toward* the observer

more waves of light are received in one second than are actually given out in that time, and consequently the wave lengths of the light received are shorter than if there were no motion. So also the motion of a luminous body *away from the earth* has the effect of increasing the length of the waves which are received from it.

Since the position of each line in the spectrum depends only on its wave length, it is evident that if a body giving a bright-line spectrum were moving toward the earth, every line in its spectrum would be shifted a little toward the blue end of the spectrum, while if it were moving away from the earth its lines would be slightly displaced toward the red. From the amount of the displacement the velocity of the source relative to the earth may readily be determined.

It has sometimes happened in examining the spectra of sun spots that when the image of a sun spot was formed on the slit of the spectroscope a part of the spot where there was a strong uprush of glowing gas has fallen on one part of the slit while a region of less disturbance has fallen on another part of it. In such a case a line in the spectrum due to this gas shows a curious twist or distortion, being displaced more toward the blue at one point than at another. Displacements pointing to a velocity of uprush of hydrogen gas as great as 400 kms. per second were observed in one instance by Young in the spectrum of solar prominences.

This method has also been used to investigate the velocity of the sun's rotation on its axis, for it is clear that in consequence of rotation one edge of the sun in its equatorial region is moving away from us while the opposite edge is moving toward us with an equal velocity.

A very interesting application of Doppler's principle is illustrated in figure 572, which shows part of the spectrum of the star Beta Aurigæ. This spectrum at times shows single lines as in the upper figure, but *once in every two days* these widen out and separate into two, "as is well shown by the calcium line near the middle of the lower diagram. The two great hydrogen lines on either side of it are too wide and too fuzzy to be separated; but the one on the left in the lower figure (which must have been printed very much more lightly than the other by some

screening process) shows that at the core it is really double, too." *

This is precisely the kind of spectrum which would be given by the light from two stars of nearly equal brightness, revolving around their common center of gravity and having their common orbit turned somewhat edgewise toward the earth. At certain times one star is coming toward us while the other is moving away from us and lines in the spectrum of the one are displaced toward the violet end of the spectrum while those in the spectrum
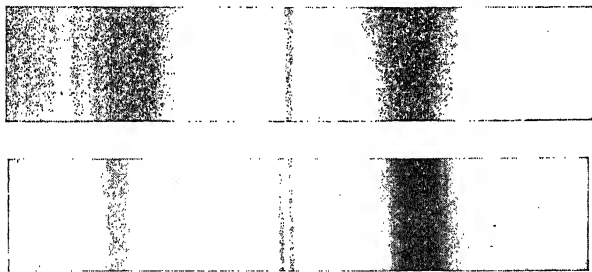
FIG. 572. Spectrum of the star Beta Aurigæ. The lines in the lower spectrum are double by Doppler's principle. Photo by Prof. E. C. Pickering

of the other are displaced toward the red, so that at such times each line is double. When the two stars are moving sidewise, say one toward the right and the other toward the left, there is no displacement of the spectra and the lines coincide. Evidently this change takes place *twice* in each complete period of revolution. It is therefore concluded that the two stars make 1 revolution in their orbit in four days.

From the amount of the displacement of the lines the maximum relative velocity of the two in the line of sight is found to be 140 miles per second. This indicates an orbital velocity of not less than 70 miles per second; and this combined with the period of revolution shows that the orbit is at least $7\frac{1}{2}$ million miles in diameter.

And now if we assume that the law of gravitation is the same for the stars as we know it to be in our solar system, and if we

* Prof. Henry Norris Russell: *Scientific American*, Sept. 3, 1910.

further assume that the two stars have equal masses, since they are equally bright, we may apply the formulas of § 164 and find that each star has twice the mass of our sun.

All of these facts have thus been obtained by the spectroscope, although the distance between the two stars is "probably at most scarcely one-fiftieth as much as that of the closest pairs which can be seen double in our greatest telescope." *

943. Motion of Stars in Line of Sight.   If the spectrum of a star shows lines which agree exactly with the spectrum lines of some known substance, such as hydrogen or iron, except that all are displaced slightly toward the red or blue, it is inferred that the lines are due to that substance and that the displacement is caused by the motion of the star either away from or toward the earth.

When it is found that most of the stars in a certain region of the heavens are approaching the earth and those in the opposite part of the sky are receding from the earth, it may be inferred that these apparent motions are probably due to the motion among the stars of our sun with its attendant earth and planets. In this way it has been concluded that our solar system is moving toward a point in the constellation Boötes about 25° north of Arcturus with a velocity that probably lies between 12 and 18 kilometers per second.

PROBLEMS

1. A spectrum line having a wave length 656.30 is displaced in consequence of the motion of the star, the apparent wave length being 656.37. Find whether the earth and star are approaching or receding from each other and with what velocity.

2. If a star is moving toward the earth with a velocity of 18 miles per sec., find the per cent of change in the wave lengths of its spectrum lines due to the motion.

COLORS OF BODIES

944. Colors of Bodies.   The colors of natural objects are due either to light waves which they themselves emit, or to their power of reflecting or absorbing the light that falls upon them

* Prof. Henry Norris Russell: *Scientific American,* Sept. 3, 1910.

from some external source. The first class includes all bodies that are self-luminous in consequence of:

(*a*) *high temperature*, as in red-hot or white-hot bodies,

(*b*) *chemical action*, as in flames,

(*c*) *electric discharge*,

(*d*) *stimulus of light from other sources*, as in fluorescence and phosphorescence.

The second class includes bodies whose colors are due either to:

(*a*) *selective absorption*, as in colored glass, pigments, and most colored bodies; or

(*b*) *selective reflection*, as in metals and bodies showing special luster.

945. Luminous Bodies. *The color of the light from any source is the average effect of its radiation upon the eye; but the particular kind of radiation which causes the effect can be determined only by analyzing the light with a spectroscope.*

For example, a yellow gas flame is found to have in its spectrum all kinds of light, but the blue and violet rays are relatively less intense than in sunlight. It is this weakness in the blue and violet which gives it a yellow color. On the other hand, the spectroscope shows that the sodium light or the yellow light obtained when a bit of common salt, previously fused, is held by a loop of platinum wire in the pale blue flame of a Bunsen burner is yellow for an entirely different reason; for the spectrum of this light consists principally of two yellow lines so close together that they appear like one line in a spectroscope of low power. The light from this flame is therefore very nearly homogeneous and appears yellow because the only kind of light present is one that excites that color sensation and no other.

946. Non-luminous Bodies. Non-luminous bodies show no color in the dark. They derive their color from the light by which they are illuminated. Let sunlight fall on a piece of colored glass or a vessel containing some strongly colored dye. The light *reflected* from the surface of the glass or solution shows no trace of color, indicating that such substances *reflect* all kinds of light equally. But light passing *through* the glass or colored solution is deeply colored and when examined by the spectroscope broad dark bands are seen in its spectrum, show-

ing· that certain constituents of sunlight have been strongly absorbed by the substance.

Thus a dark blue cobalt glass transmits green, blue, and violet, but absorbs strongly yellow and orange and most of the red. The curve in figure 573 represents by its height the intensity, in different parts of the spectrum, of light which has passed through a certain thickness of this kind of glass. Such absorption is called *selective*.

Red – Orange –Yellow – Green – Blue– Violet

FIG. 573. Spectrum of light transmitted by blue cobalt glass

947. Spectrophotometer. An instrument in which the spectra from two sources are formed side by side and with appliances so that the relative intensities of the two spectra can be determined for each point in the spectrum is known as a *spectrophotometer*.

Such a curve as that shown in figure 573 is obtained by means of an instrument of this kind, the spectrum of direct sunlight being compared with that of sunlight which has passed through the colored substance.

948. Absorption and Color of Powders. The most common cause of the color of bodies is *absorption*. A crystal of copper sulphate when seen by ordinary daylight appears blue because light coming to the eye through the crystal has lost the red and yellow rays by absorption. The light received by the eye is, however, not a pure blue, but is diluted with white light reflected from its surface. If the crystal is broken into smaller fragments the thickness of crystal through which the transmitted light passes before meeting a reflecting surface is smaller and there is accordingly less absorption and the blue color ·is not so marked. If the crystal is finely pulverized, the dry powder appears a pale whitish-blue, for light can penetrate only to an extremely small depth before being reflected and scattered by the surfaces of the tiny fragments.

From the above considerations it is evident that *all clear colorless substances, such as ice, glass, Iceland spar, etc., must make white powders*, since they reflect and scatter the light from innumerable minute surfaces but absorb scarcely any of the visible rays.

The light reflected to the eye by such a powder is, therefore, of the same quality as that which falls upon it, and when illuminated by white light it appears white.

949. Effect of Illumination. Except in case of self-luminous bodies, *the color of a substance depends on the light by which it is illuminated.* When a piece of red paper is held in the red of a bright spectrum it appears bright red, but when held in the yellow, green, or blue parts of the spectrum it appears black, for it can reflect red rays, but it absorbs the yellow, green and blue. So a blue paper may reflect the violet, blue and green, but will appear black in the red, orange, or yellow parts of the spectrum, while a white paper reflects whatever color falls upon it.

Two kinds of light that appear very much alike in color may yet have very different effects on the colors of bodies. For example, an ordinary gas flame gives out a yellowish light not very unlike the yellow *sodium flame* in appearance, and yet bright-colored objects or pieces of paper are seen in their various colors when illuminated by the gas flame, but all appear of one color, either brighter or darker yellow or black, when illuminated with the sodium flame. This is because the ordinary flame gives out all kinds of light waves from red to violet, while the sodium light is nearly *homogeneous*. The peculiar ghastly appearance of persons illuminated by a salted alcohol flame is due to this cause.

The difference between colors seen by daylight and gaslight is because light from the blue end of the spectrum is relatively far more intense in the former than in the latter.

950. Matched Colors. Two colors that appear alike by daylight may yet be due to very different kinds of light. When an object appears yellow it does not follow that it reflects only rays from the yellow part of the spectrum. It means simply that the stimulus given to the retina by the various kinds of rays coming from the object excites the same sensation as the yellow light of the spectrum. What particular waves cause the color can be determined only by dispersing the light and examining its spectrum.

Consequently two colors which match perfectly when seen by daylight may differ very much when illuminated by some artificial light.

951. Mixed Pigments. *When paints are mixed the resulting*

color is not a mixture of the colors that each would give separately, but is due to the double absorption which light suffers in the mixture. For instance, a solution of gamboge yellow absorbs all rays but yellow and green, while a solution of Prussian blue absorbs all but blue and green, a mixture of the two will therefore transmit only the green.

952. Mixing Colors. To find the color that will be produced by a mixture or blending of colored lights, the lights to be mixed may be made to illuminate simultaneously a white screen, or the color wheel may be employed. In this apparatus discs of colored paper, each slit from center to edge, and fitted together so as to expose a sector of each color, are mounted on a spindle and rapidly rotated. If the speed of rotation is sufficient the disc appears of a uniform color which is the mixture of the different colors used. The proportional amounts of the several colors depends on the widths of the exposed sectors and may be changed by slipping the discs on each other. A larger and smaller disc may be mounted on the same spindle for comparison.
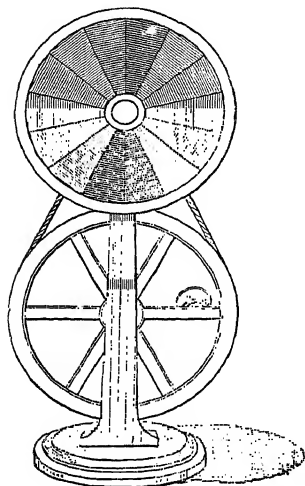
FIG. 574.   Newton's color disc

The effect in this case depends on the persistence of the sensation for a very short time after the stimulus to the retina has ceased. The various colors give their stimuli in such rapid succession that the effect is a blended sensation.

Newton found that a color disc painted in sectors to imitate the colors of the spectrum appeared grayish-white when rapidly rotated and could be matched with a black disc having a white sector, black being used to diminish the intensity of the white.

Complementary Colors. Two colors which when combined produce white, are said to be *complementary*. By means of the color disc it is found that blue and yellow of the proper tints and intensities will make white, also green and red may be complementary.

953. Metallic Luster. Metals owe their peculiar luster to their intense reflecting power. Polished silver reflects 90 per cent of the light that falls upon it, while glass at perpendicular incidence reflects less than 5 per cent.

Sunlight reflected from red or blue glass remains white, but when reflected from gold-leaf it is yellow. This shows that the reflection of light in case of some metals is *selective*, some kinds of light being more strongly reflected than others. It is to this property that the colors of metals are due.

The light transmitted through a thin film of gold-leaf is not yellow, but green. The yellow light which is reflected is that also for which the absorbing power of the metal is greatest.

Some non-metallic substances also have the power of reflecting light like metals as is seen in the bronzy luster of aniline ink and in crystals of permanganate of potash. Such substances show strong selective absorption and anomalous dispersion.

954. Fluorescence. When a strong beam of sunlight or light from the electric arc is sent through a block of glass colored with oxide of uranium the transmitted light is yellowish, showing that there has been absorption of the shorter wave lengths; but besides this the whole block of glass is seen to glow with a greenish light which seems to come from each point in the glass itself, making the whole block seem turbid and milky. This is called *fluorescence*, for the phenomenon is strongly marked in fluorspar.

The subject was first carefully investigated by Sir George Stokes, who showed that fluorescence is really a kind of radiation from the molecules of the substance under the stimulus of the absorbed light. Light is absorbed by the block of glass, and the energy of the absorbed waves, instead of appearing simply as heat, produces an excitation of the atoms, and the energy shifts during readjustment produce light waves from these atoms, some of the same frequency as that absorbed but also some of longer wave length. Stokes announced the law that *waves of fluorescent light cannot be shorter than the absorbed waves to which they are due.* This law is a necessary consequence of the absorption and emission of light waves in quanta as explained briefly in the last chapter of this book.

It will be noticed that the block of glass fluoresces most strongly near the side where the incident beam enters, for as the beam penetrates into the block it loses by absorption the very rays which are effective in causing fluorescence.

The interposition of a piece of red glass in the path of the light

cuts off all fluorescence, while a blue cobalt glass scarcely weakens it at all, showing that the effect is due to the shorter wave lengths which are transmitted by the blue glass but suppressed by the red.

Many substances show fluorescence, among others almost all mineral oils, especially the thick heavy oils, and crude petroleum, and even refined kerosene oil shows a delicate blue fluorescence in strong light. Some of the anilin substances are extremely fluorescent, notably fluorescein and eosin. Sulphate of quinine fluoresces a delicate blue as does also æsculin obtained from crushed horse-chestnut bark.

A white card covered with a thick paste of sulphate of quinine moistened with dilute sulphuric acid will fluoresce strongly in the invisible rays of the spectrum beyond the violet, the so-called ultraviolet region.

955. Phosphorescence. When fluorescence persists after the illumination ceases the substance is said to be *phosphorescent.*

By a special contrivance, called a phosphoroscope, Becquerel found that many substances, including paper, bone, and ivory, not usually known as phosphorescent, glow for a fraction of a second after the incident beam is cut off.

The sulphides of calcium, barium, and strontium are strongly phosphorescent and the color of the phosphorescent light is greatly influenced by the presence of slight impurities.

This kind of phosphorescence may be called *physical* to distinguish it from the glow of decaying vegetables, of fire-fly and glow-worm, and of phosphorus itself, in which the light seems to be due to *chemical* changes.

956. Theory of Color Sensation. The Young-Helmholtz theory of color sensation proposed by Thomas Young and modified by Helmholtz assumes that light falling on any point in the central region of the retina where it is sensitive to colors, excites in general three primary color sensations, red, green, and blue, the resulting color sensation depending on the relative intensities of these three primary sensations.

The sensation of red is found to be excited more or less by all wave lengths in the visible spectrum, but most strongly by the long waves, as shown in the left-hand curve of figure 575. So that if a person possessing only the red color sensibility and lack-

ing those of green and blue were to look at a bright spectrum it would appear to him red from one end to the other, but brightest where the wave lengths are long as shown in the curve marked *red*. So, too, the curve marked *green* may be taken as exhibiting the relative intensity of the green sensation excited by different wave lengths of light, while the third curve shows how the sensation of *blue* varies with the wave length.

In the normal eye, possessing all three sensibilities, a given wave length of light excites all three sensations, the red predominating in case of long waves, green when the waves are
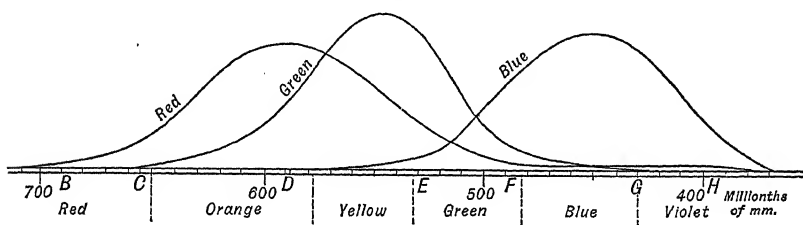
Fig. 575.   Curves from Abney, showing variation of color sensation with wave length, according to the Young-Helmholtz theory

shorter, and blue when they are shorter still, the intensity of each sensation being proportional to the height of its curve at the point corresponding to the given wave length.

The sensation of white results when all three of the primary sensations are equally excited.

What the three primary color sensations are, can be determined only by the study of color-blind individuals. By such a study Koenig finds that the primary sensations are the red, green, and blue found in the spectrum at wave lengths, 671, 505, and 470$\mu\mu$, respectively. By combining these three colors in proper relative intensities any color of the spectrum may be produced.

Helmholtz assumed that there were three kinds of nerve termini in the retina corresponding to the three primary sensations of color, while Hering supposes certain substances in the retina whose transformations under the influence of light give rise to the various primary sensations. For a further discussion of theories of color vision the reader may consult *A First Book in Psychology*, by Calkins, or the article "Vision" in Baldwin's *Dictionary of Philosophy and Psychology*.

INTERFERENCE OF LIGHT

957. Introduction. Up to this point in our study the theory that light is a wave motion has been supported by the fact that the velocity of light is the same as that of electric waves and by the simple explanation which that theory affords of the phenomena of reflection and refraction. But we have not yet found any direct evidence of the existence in a beam of light of a regular periodic oscillatory motion such as is characteristic of all kinds of waves. We now come to some phenomena which point unmistakably to just such a periodicity.

958. Interference of Waves. Perhaps the most distinctive evidence of wave motion is afforded by the phenomena of *interference*.

When two trains of waves come together having the same wave length and amplitude and traveling in nearly the same direction, there will be found *points of rest or of very slight motion where the two systems of waves are in opposite phases and neutralize each other*, and other points where the waves coming together in *the same phase cause an amplitude of motion equal to the sum of the amplitudes of the component waves.*

The interference of water waves and sound waves has already been discussed (§ 332).

959. Young's Experiment. The interference of light waves was first shown by Thomas Young in 1801 by the method illus-
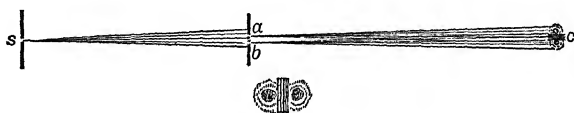
FIG. 576. Young's experiment showing interference

trated in the diagram. In the path of a beam of sunlight shining through a minute pinhole at S, is placed a screen of tinfoil having two very small holes a and b close together. If light is now allowed to pass through only one of the openings, a round bright spot surrounded by faint dark and bright rings is formed on a screen at C. But if light passes through *both* openings, there is seen between the two bright spots and at right angles to their line of centers, a series of bright and dark bands, as shown in the lower part of figure 576.

The explanation of these bands will be understood by the aid of figure 577. Waves from S set up waves at a and b which start out simultaneously in the same phase, the two sets of waves spreading out in the medium beyond, one set from a and one from b, as shown in the diagram. The central point c is equidistant from a to b, so that waves leaving a and b at the same instant meet at c in the same phase, reënforcing each other and making c a bright spot. But d is *a half wave length farther*
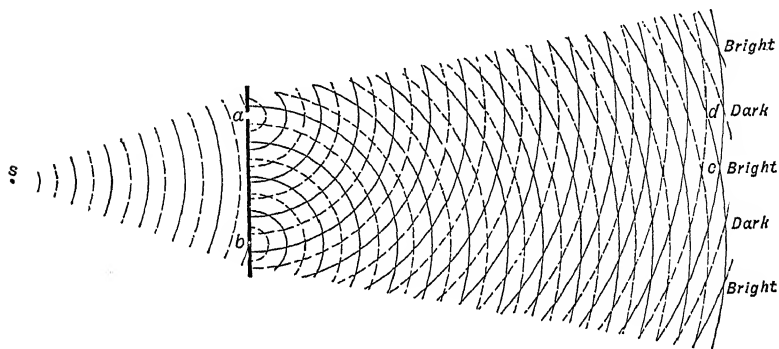
FIG. 577.   Diagram of interference of waves.   Young's experiment

from b than from a, and consequently waves from a and b reach there in opposite phases and neutralize each other, making d a dark spot. In this way those points on the screen which are equidistant from a and b, or which are one, two, or more whole wave lengths farther from one opening than from the other will be bright, while points which are farther from one opening than the other by $\frac{1}{2}$ or $1\frac{1}{2}$ or $2\frac{1}{2}$, etc., wave lengths, will be dark.

960. Fresnel's Interference Experiment. In order to show that the above explanation of the dark bands obtained by Young was correct and that they were really due to interference of waves, Fresnel devised a most ingenious modification of the experiment, by which he avoided any disturbance of the light that might be imagined to result from its passing through the small openings a and b.

Waves of light from a narrow slit shown in section at S (Fig. 578) fell on two mirrors $M'M''$ inclined to each other at a small angle so that light after reflection from M' diverged as if

from S', while that reflected from M'' came as if from S''.   In this way two trains of waves were produced which gave bright bands at c, f, and g, where the waves of the two sets were in the
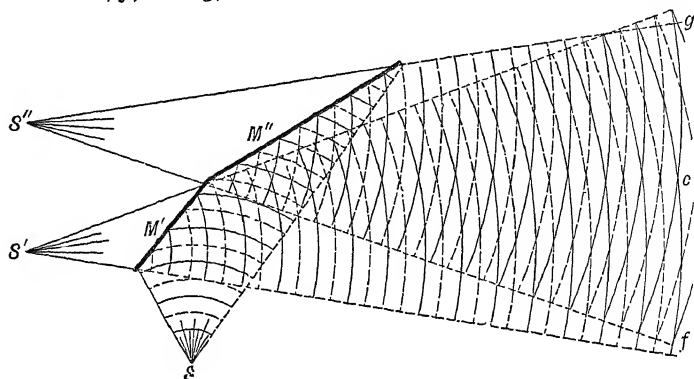
FIG. 578.   Fresnel's interference experiment with mirrors

same phase, and intermediate dark bands, just as in Young's experiment.

961. Newton's Rings and Colors of Thin Films.   When a lens having a convex surface of *very* slight curvature is placed in contact with a flat glass plate a thin film of air is enclosed between the two plates which increases in thickness from the central point of contact outward.   If this film is examined in white light, holding the eye so as to receive light reflected from its surface, a number of colored rings are seen surrounding the central point of contact of the lens and plate, each colored ring corresponding to a definite thickness of the air film.
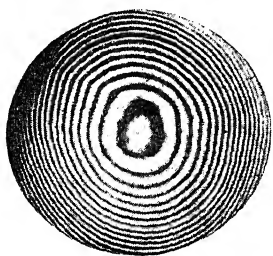
FIG. 579.   Newton's rings in sodium light

If illuminated with light of one wave length, as sodium light, the whole surface of the film is seen to be covered with alternate dark and bright rings.   When red light is used each ring is larger than the corresponding ring in case of blue light.

The colored rings observed in white light are due to the superposition of the sets of rings of different sizes due to different wave lengths of light.

These bands are known as Newton's rings, for the experiment was devised by him to determine the thickness of film corresponding to a given color; for if the curvature of the lens surface is known, it is easy to calculate the thickness of the air film for a ring of any given radius. The following table shows some results:

THICKNESS OF FILM IN NEWTON'S RINGS

Red light		*Blue light*	
1st ring,	0.00017 mm.	1st ring,	0.00012 mm.
2d ring,	0.00051 mm.	2d ring,	0.00036 mm.
3d ring,	0.00085 mm.	3d ring,	0.00060 mm.
Difference,	0.00034 mm.	Difference,	0.00024 mm.

In a similar way, if two perfectly flat pieces of glass are laid one upon the other, touching at one edge and separated at the other by a thin strip of tinfoil, the wedge-shaped film of air formed between them shows alternate dark and bright bands in homogeneous light. These bands are straight and parallel to the edge of the wedge if the plates are flat, and the straightness of the bands affords a sensitive test of the flatness of the plates.

The English physicist, Thomas Young, first showed that Newton's rings could be explained easily by the interference of light waves as follows:

Let a train of waves advance upon a transparent film, and for simplicity suppose the incidence to be nearly perpendicular, as shown by the arrow (Fig. 580). Each wave on meeting the surface is partly reflected and partly refracted. An advancing wave meeting the surface at *a* is in part reflected along *ac*.

FIG. 580. Interference in reflection from a film

Another part of the same wave passes into the film at *d*, meets the second surface at *b*, and is in part reflected to *a*, where it emerges along the direction *ac*. This portion of the wave will have had to *cross the film twice* and will therefore have fallen behind the part which was reflected directly at *a*, so that if the thickness of the film is one-quarter of the

wave length of the light waves in the film, waves from d will reach a one-half wave length behind the corresponding waves reflected at a, and may therefore be expected to interfere, causing the film to appear dark at a to an observer looking in the direction ca. At points where the thickness of the film is one-half a wave length each wave from d reflected at b will be a whole wave length behind the corresponding wave reflected at a and may therefore be expected to reënforce the next succeeding wave, and so the film should appear bright at such points. But experiment shows that *exactly the reverse* is true, for the central point of contact in Newton's rings is dark by reflected light instead of being bright.

Thomas Young showed that this discrepancy is due to a *change of phase which takes place in the very act of reflection;* for at one surface of the film, waves in a less refracting medium are reflected where they meet the more refracting one, while at the other surface, waves in the more refracting medium are reflected on meeting the less refracting one. These two reflections are opposite in kind, just as reflection in a stopped organ pipe is opposite to that in an open pipe (§ 349), and one changes the phase of the reflected light while the other does not.

This opposition of phase brought about in reflection exactly reverses the conclusions reached above where merely the effect of the thickness of the film was considered, and consequently *those parts of the film where the thickness is an odd number of quarter wave lengths appear bright by reflected light.*

. To test whether this explanation was correct, Dr. Young reflected light from a thin film of oil of sassafras between a lens of crown glass on one side, and a flint glass plate on the other. The index of refraction of the oil of sassafras is more than that of crown glass, but less than that of flint glass, so that the change in phase due to reflection was the same at each surface. It was found in this case that *the central spot where the film was thinnest was bright* by reflected light, while dark bands were observed where the thickness of the film was one-quarter of a wave length, three-quarters of a wave length, etc., thus confirming Young's idea as to the cause of the reversal.

The thickness of the film for red and blue rings given on page 679 therefore leads to the conclusion that the average wave length of the red light used was about 0.00068 mm., while that of the blue light was 0.00048 mm.

The colors of soap bubbles and of thin films of oil or turpentine on water are also explained in the same way. For instance, where the thickness of the film is equal to a half wave length of red light it will be nearly equal to three-quarters of a wave length of blue light. The former will therefore be destroyed by interference, while the latter will be reflected and the film will appear blue.

A soap bubble does not show color unless it is very thin; for when the thickness of the film is, say, 0.0028 mm., it will be equal to 4 wave lengths of extreme red light and 6 wave lengths of extreme blue; these waves will be absent from the reflected beam because of interference, and also the waves whose lengths are such that in the thickness of the film there are included just 4½, 5, and 5½ wave lengths, respectively; while light having wave lengths intermediate to these will be reflected from the film. The film in such a case appears white because the reflected light contains so many different wave lengths that the average effect is white.

The intensely black spots seen in thin soap-bubble films by reflected light have been found by Reinold and Rücker to have a thickness of only about one-fiftieth of the wave length of sodium light, and appear black for the same reason that the central spot is black in Newton's rings.

962. Interference with Great Difference of Path. In case of thin films the interfering waves differ in path by only a few wave lengths. There are some cases, however, in which interference has been obtained when the difference in path is very great.

A useful form of *interferometer*, as it is called, is that of Michelson, a diagram of which is shown in the figure.

A plate of glass A, having plane parallel surfaces, is mounted in front of the mirror M in an oblique position so that light from the source L on meeting the second surface at S is partly reflected to the mirror M and in part transmitted to a second mirror N which is at right angles to the first. The eye at E will therefore receive light from S which has been reflected at M, and also light from the same point which has been reflected at N, and when the adjustment of the instrument is correct, these two rays will interfere when they come together if the light, in going from S to M and back, has to pass over a distance which differs by an odd number of half wave lengths from the distance from S to N and back again.

FIG. 581. Michelson's interferometer

In order that the interference may be complete, the reflecting surface

at S has a very thin coating of silver, just sufficient to make the reflected and transmitted beams of light equally intense.

The mirror M is usually mounted so that its distance from S can be varied by means of a micrometer screw. The plate of glass B of the same thickness as A is mounted parallel to it, so that waves reflected at N have to pass through the same thickness of glass as those reflected at M. The interference bands in this case are circles, which expand and are succeeded by others, as the mirror M is moved away from S. A motion of M through one-half of a wave length will cause a shift in the position of the interference bands equal to the distance between two successive bands.

Using this method and employing light of one wave length from a Plücker tube (§ 937) containing mercury vapor, Michelson obtained interference bands when one path was longer than the other by 540,000 wave lengths.

By counting the bands that pass when the mirror M is moved backward a certain distance by the screw, the number of wave lengths of light contained in that distance may be exactly determined. By a very ingenious extension of this method, which the student will find described in detail by Professor Michelson in *Light Waves and Their Uses*,* the length of the standard meter was determined by him in terms of wave lengths of light. The results obtained were

$$1 \text{ meter} = \begin{cases} 1,553,163.5 \text{ waves of the red radiation from cadmium.} \\ 1,966,249.7 \text{ waves of the green radiation from cadmium.} \\ 2,083,372.1 \text{ waves of the blue radiation from cadmium.} \end{cases}$$

all in air at 15° C. and normal pressure.

Of these results Michelson says: "It is worth noting that the fractions of a wave are important, because, while the absolute accuracy of this measurement may be roughly stated as about one part in two millions, the relative accuracy is much greater, and is probably about one part in twenty millions."

REFERENCES

A. A. MICHELSON: *Light Waves and Their Uses.*
EDWIN EDSER: *Light for Students.*

DIFFRACTION

Diffraction Bands Around Shadows. The observation hadows suggests that light is propagated in straight lines. : form which the shadow of an obstacle would have if this e the case is called the *geometrical shadow;* it is the projection he obstacle upon the screen by straight lines radiating from luminous source of a center.

* Univ. of Chicago Press.

Ordinary shadows are blurred at the edges because the angular magnitude of the source causes a penumbra. Hence to make an accurate comparison of a real shadow with the geometrical shadow the source of light should be a mere point. But when the experiment is tried, as, for example, when we examine closely the shadows cast by a moderately distant arc lamp, instead of finding a clear-cut boundary, *the entire edge of the shadow is observed to be surrounded by a series of alternate dark and bright bands, parallel to the edge,* very distinct next to the shadow and gradually fading out into the fully illuminated region.

These bands were known at the time of Newton and were called *diffraction bands* or *fringes*, because to explain them on the emission theory it was supposed that the luminous corpuscles were bent aside from their straight course as they shot by the edge of the obstacle.

In the year 1816 a young French artillery officer, Joseph Fresnel, then less than thirty years of age, presented to the French Academy a memoir which marked an epoch in the science

FIG. 582. Diffraction bands at the edge of a shadow. The vertical line is the edge of the *geometrical* shadow

of optics, for in it he showed that the varied phenomena of diffraction are readily explained in every detail by the interference of light waves taken in connection with Huygens' principle, and without recourse to any additional hypothesis.

964. Huygens' Principle. Let there be a train of waves advancing in the direction OP whose crests are represented by the parallel lines on the left of figure 583, and let AB be a row of particles parallel to the wave front; then as the waves sweep by AB the particles are all set vibrating simultaneously and in the same phase. Now, each of these vibrating particles may be considered as a center of disturbance from which spherical waves spread out into the region beyond. Thus, if we choose, *we may consider the vibration that is produced at the point P as due to*

the combined effect of all these little elementary waves or wavelets whose centers lie in the line AB, just as though the line of particles AB was the actual source from which waves spread out. This is known as Huygens' principle. Figure 186b shows an interesting illustration of Huygens' principle for sound waves.

965. Diffraction by a Narrow Slit. If sunlight shining through a narrow slit falls on a second narrow slit parallel with the first,
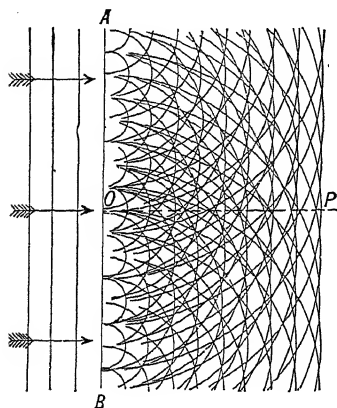
Fig. 583. Huygens' principle

there will be seen on a white screen held back of it a central bright band and on each side alternate bright and dark bands, which widen out when the second slit is made narrower. The wave theory affords a simple explanation; for let S be the slit (looked down upon endwise) which is *very* narrow compared with its distance from the screen at b_1 (Fig. 584), and let AB represent the greatly magnified cross section of the slit in the plane of the paper. Then the "ether particles" lying in ACB are kept in vibration by the successive waves passing through the slit, and by Huygens' principle these particles may be considered as the centers of wavelets which spread out in all directions and
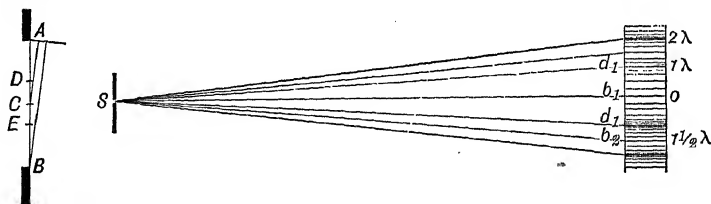
Fig. 584. Diffraction through narrow slit perpendicular to the plane of the paper

produce the effects which are observed. Now, on account of the extreme narrowness of the slit, b_1 is practically equally distant from all points along the line AB, and therefore the wavelets starting simultaneously at all points along AB reach b_1 in the same phase and so reënforce each other and make it a bright spot.

Just below b_1 there will be a point d_1 which is a *whole wave length* farther from A than from B. Then d_1 is a *half* wave length farther from C than from B, and *for every point between B and C there is another between C and A which is just a half wave length farther from d_1. Therefore the wavelets going to d_1 from one-half of the slit will be exactly neutralized by wavelets from the other half, and d_1 will therefore be a dark spot in consequence of this interference.* In the same way the dark spot d_1 above b_1 is explained.

But a little beyond d_1 there will be a point b_2 which is $1\frac{1}{2}$ wave lengths farther from A than B. In that case the wave front in the slit may be conceived as divided into three equal parts AD, DE, and EB, such that wavelets coming to b_2 from AD have a half wave length farther to travel than from the corresponding point in DE. Therefore waves from these segments interfere at b_2 while waves from the third segment will be effective and make b_2 a bright spot, though *much less bright than b_1*, since only one-third of the width of the slit is effective. Of course DE might be regarded as opposing EB, and in that case AD is effective.

The same reasoning shows that there will be a dark spot where the difference in path from A and B amounts to 2 wave lengths and again a bright spot where the difference amounts to $2\frac{1}{2}$ wave lengths. There will therefore be a series of alternate dark and bright spots on each side of b_1 as experiment shows.

If the slit is made narrower the line AB is shorter, and consequently the point d_1, which is one wave length farther from A than from B is farther away from b_1 than before. Therefore the bands spread out as the slit is made narrower, and if it had a *width* of *only one wave length* or less, light would go out from it in every direction, though the intensity would be less in oblique directions in consequence of partial interference.

966. Shadow of a Circular Obstacle. When Fresnel's memoir was presented to the French Academy it was objected by Poisson that if his views were correct there should be a bright spot in the center of the shadow cast by a circular disc. Fresnel at once acknowledged the justice of the criticism and, making the experiment, *found the bright spot,* thus obtaining a triumph for the new theory.

The experiment may be made by fastening to a piece of plate glass a bicycle ball about $\frac{1}{4}$ inch in diameter and observing its shadow as cast by a distant arc light at a distance of 8 or 10 ft. back of the obstacle; the central bright spot may readily be seen either by receiving the shadow on a card or by looking toward the object and viewing the shadow directly with a small pocket magnifier. Or the shadow may be received upon a sensitive film and photographed.

The central spot is bright because it is equally distant from every part of the edge of the obstacle, and therefore wavelets
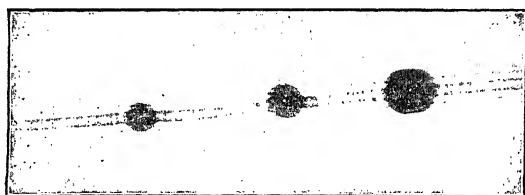
Fig. 585.   Shadow cast by small balls fastened to a fine wire

coming from points just outside the edge around its whole circumference come together *in the same phase* at that point.

Similarly a bright line is found in the center of the shadow of a wire, since the central line is equidistant from the two edges and waves coming around the wire on both sides reach the central line in the same phase and therefore reënforce each other. (See figure 585.)

The student should observe through a pocket magnifier the diffraction bands formed by the wires of a mosquito netting or screen of thin silk,
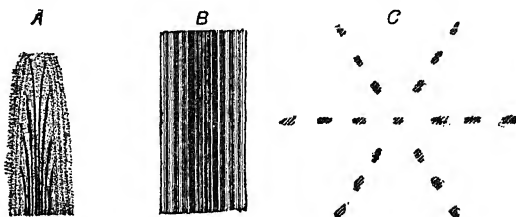
Fig. 586

standing a few feet from the screen and looking through it toward a distant arc light.

967. Miscellaneous Diffraction Phenomena. In ·figure 586 are shown at *A* the diffraction bands in the shadow of the pointed

end of a needle. It will be observed that there is a central bright band, broadest near the very point, while where the needle is thicker many fine interference bands are seen in the shadow. This is shown in *B* which is the shadow of a somewhat thicker wire showing the many fine bands due to the interference of the waves coming around the two sides of the wire. At *C* is shown the diffraction pattern which may be seen by looking through the cloth of a silk umbrella toward an electric arc lamp.

A small round obstacle gives rise to a series of diffraction rings, and where the rings due to a great number of fine particles are all of the same size and are superposed the effect may be very intense. This is the explanation of the *coronas* seen so often around the moon. They are brightest when the light from the moon comes through a region full of minute water particles *nearly uniform in size*. These coronal rings are larger the smaller the particles that cause them, and the average diameter of the water drops can be immediately calculated from the angular radius of the rings.

Beautiful coronas may be seen on looking at an electric light or gas flame through a piece of glass coated with *lycopodium* powder, which is made up of minute discs of nearly uniform size. First breathe upon the glass, then pour some of the powder upon it and shake off the loose dust.

968. Diffraction in Case of a Lens. In our study of lenses we saw that a lens transforms a wave of light coming from a distant point into a
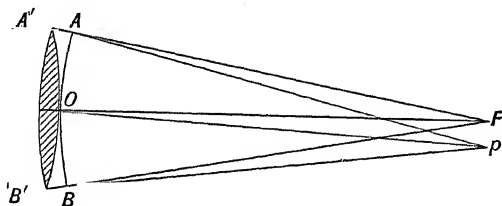
FIG. 587

concave spherical wave which has its center at the focus toward which it converges. Thus the wave *A'B'* becomes concave, as at *AB*, and *if the latter is perfectly spherical the lens is perfect. The geometrical theory of optics would lead us to infer that in that case the light would all converge rigorously to the point F, but the wave theory shows that this cannot be so.*

To determine the effect at *F* of the wave *AOB* we must again have recourse to Huygens' principle and consider the resultant effect as due to wavelets having their centers in the concave surface *AOB*. Clearly all will reach *F* in the same phase, since it is equidistant from all, hence *F* must be

a point of maximum brightness. A little below F there must be some point p which is on the average a half-wave length farther off from the upper half of the surface AOB than it is from the lower half. At that point waves from one half of the surface will interfere with those from the other half and produce complete darkness. But between p and F the interference is only partial and consequently the light intensity must shade off from F to p.
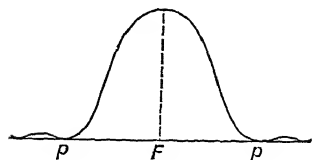
Since the light is symmetrical about OF, there must be a little *spot* of light formed at the focus, having the distance Fp as its radius. The curve in figure 588 shows by its height how the intensity of the light in the focal spot varies from F to p.

FIG. 588

969. Resolving Power of Optical Instruments. The fact that the focal spot has an appreciable size has a most important bearing on the *resolving power* of optical instruments, for when a lens forms an image of any object each *point* in the object is represented by a little *spot* in the image, and the sharpness of definition in the image depends on the *smallness* of the focal spots.

Now, it may be proved that the *effective diameter* of the focal spot, or *diffraction image of a point,* as it is called, is equal to $\dfrac{\lambda F}{D}$, where λ is the wave length of light, D is the diameter of the lens, and F is its focal length. Hence for a given focal length the focal spot will be smaller the larger the lens.

The *angular diameter* of the focal spot is $\dfrac{\lambda}{D}$ which is equal to 4.5″ of arc when $D = 1$ inch. Therefore a telescope having a perfect object-glass 1 inch in diameter will be just capable of resolving a double star whose components are 4.5″ apart. For in that case the star images formed in the telescope will be two spots of light just touching each other. If the object-glass is 2 in. in diameter it may then be capable of resolving stars only 2.3″ apart. Evidently no magnification by the eye-piece will increase the resolving power as it will simply show two larger spots of light touching each other instead of two smaller ones. Helmholtz has shown that in consequence of the size of the focal spot it is impossible to have a microscope that will enable the eye to distinguish separate lines which are less than $\frac{1}{135,000}$ of an inch apart, and even this limit can be reached only by oil-immersion lenses.

970. Diffraction Grating. One of the most useful instruments for the formation of spectra and for the measurement of the length of light waves is the *diffraction grating*, so called because the first gratings made by Fraunhofer were veritable gratings made of fine parallel wires spaced at equal intervals. More accurate gratings are made by ruling with a diamond on

glass or on a polished mirror surface of speculum metal an immense number of parallel equidistant straight lines, and copies or replicas of these ruled gratings are made by photography or by direct impression on a plate of celluloid.

The effect of such a grating, of the transparent sort, is shown in the figure below. At S is placed a narrow slit upon which is concentrated a beam of sunlight, the slit is supposed perpendicular to the plane of the paper so that its section is shown at S. In front of the slit is placed a lens L which forms a sharply defined image of the slit on the distant screen at O. If the grating is now interposed as shown, with its bars or rulings parallel with the slit, there are seen upon the screen several spectra on each side of the central image, which are said to be of the first,
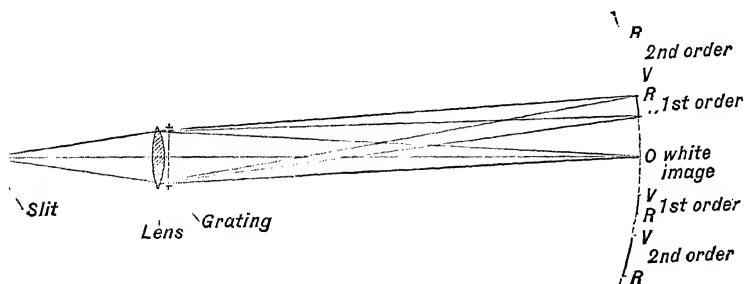
Fig. 589. Spectra formed by diffraction grating

second, or third order, etc., according to their distances from the center. These spectra have their violet ends toward the center and their lengths are nearly proportional to the numbers expressing their orders.

If the grating has only two or three bars to the millimeter the spectra will be very narrow, forming a group of bright bands on each side of the central image. But as the rulings are made closer together the spectra are longer and more spread out.

Very perfect gratings were made by Professor Rowland, of Baltimore, on a ruling engine devised by him. In many of these gratings 14,438 lines are ruled to the inch, or about 568 lines per millimeter.

971. How Gratings Produce Spectra. Let the grating consist of a set of opaque bars, which are represented in cross section, greatly magnified, by the heavy lines in figure 590.

When a series of flat waves comes from the left, as shown by the arrows, the ether particles in the openings *abc*, etc., are simultaneously set in vibration, and by Huygens' principle each particle is a center from which wavelets spread out in all directions into the region beyond.

Now, if a convergent lens is placed in front of the grating as shown at L, a flat wave parallel with the grating will be converted by the lens into a concave wave converging upon its principal focus at O, the lens retarding the middle portion of the wave more than the edges, so that all parts reach O at the same instant. Therefore wavelets starting simultaneously from all the grating
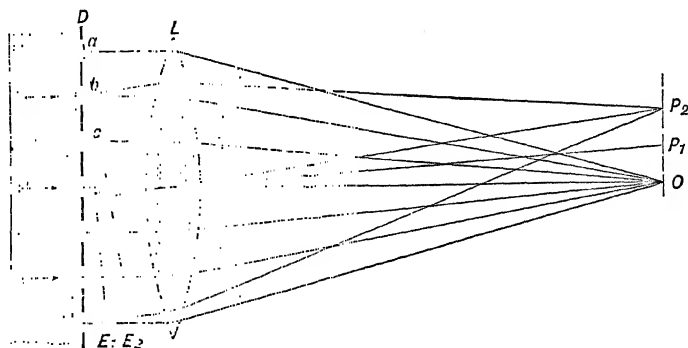
FIG. 590

openings will by the effect of the lens reach O at the same time and in the same phase. *The point O will therefore be bright whatever may be the wave length of the light.*

In the same way by the effect of the lens an oblique wave parallel to DE_1 is brought to focus at P_1 on the line through the center of the lens and perpendicular to DE_1. Therefore, in consequence of the lens it takes light equally long to reach P_1 from any point whatever on DE_1, and consequently wavelets from the grating that agree in phase on reaching DE_1, will also agree in phase at P_1.

Suppose that DE_1 is drawn through the edge of one grating space in such a direction that it is distant exactly one wave length from the corresponding edge of the next grating space, as shown in figure 591 where *fe* is supposed just equal to a wave length. Then a wavelet starting from a point in the opening *a*

will reach DE_1 at the same instant as the wavelet which started from the corresponding point in b just *one complete period before,* and *the two wavelets will therefore reach DE_1 in the same phase.* So also the wavelet reaching DE_1 from the corresponding point in c will agree in phase with those from a and b, and thus as wavelets from all the openings reach DE_1 in the same phase they
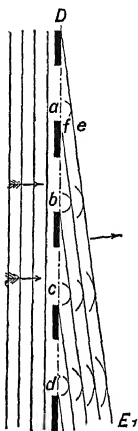
FIG. 591
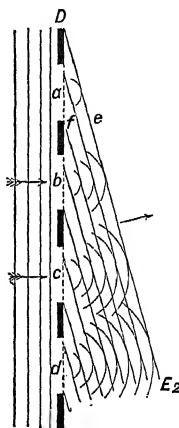
FIG. 592

will agree in phase at P_1, which will therefore be bright, and may be called the *first order* bright spot.

Let us now consider a line DE_2 (Fig. 592) so oblique that fe is equal to two whole wave lengths. Then again wavelets from corresponding points in all the openings will agree in phase on reaching DE_2 and consequently the point P_2 to which they are converged by the lens is a bright point. Thus on each side of the central spot at O there will be bright spots of the first, second, etc., orders.

If the light were homogeneous or all of one wave length there would be only these bright spots, all of the same color; for instance, if the grating is illuminated with sodium light there will appear a central yellow band with narrow yellow bands on each side somewhat as shown in the upper row in figure 593. With a source giving out longer waves, as of red light, the bands would be farther apart; for the distance fe in figure 591 would be greater and hence the line DE_1 would be more inclined, making

the point P_1 farther from the center. While if the waves were shorter the bands would be closer together, as shown at V_1V_2 in the third row of figure 593. Consequently when *white* light shines upon the grating, having all wave lengths present, each wave length produces a bright band at the appropriate distance from the center, and therefore there results the spectra of the different orders represented in the lower part of the figure, the violet end of each order being toward the center, showing

FIG. 593. Formation of diffraction spectra

that wave lengths increase from the violet toward the red end of the spectrum.

672. Effect of Removing Grating. If the grating is removed the side spectra all vanish, leaving only the central image at O.

FIG. 594

For the wavelets which were cut out by the bars of the grating now interfere with the wavelets which formed the side spectra. This is easily seen by a consideration of figure 594. Let ac be drawn from the edge of one grating space so that bc, its distance from the corresponding edge of the next space, is one wave length, then bc will be the direction in which the first order spectrum is formed. Imagine the grating bar between a and b removed so that light may now proceed from all points between a and b. Let e be a point half-way between a and b, then ef is one-half a wave length, and corresponding to any point h between a and e there is a point g between e and b which is just a half wave length farther from ac. Waves therefore which start in the same phase from h and g must reach ac in opposite phases, and consequently light going out in the direction bc from points between e and b will be exactly interfered with and neutralized by light from points between e and a, and *there will therefore be no first order spectrum.* In a similar way

it may be shown that if the grating bars are removed there will be no side spectra of any order.

973. Resolving Power of Grating. A grating should have a large number of bars and spaces for two reasons. First, the brightness of the diffraction spectra will be greater the larger the number of grating spaces. And, second, the power of a grating to give a sharply defined spectrum is proportional to the total number of grating spaces, other things being equal. For let AB, figure 595, represent a grating of 1000 spaces, and let AD be so drawn that its distance from the first grating space next to A is one wave length, from the second space its distance is 2 wave lengths, etc., from the 500th space at C its distance is 500 wave lengths represented by CE, and from the 1000th space at B its distance BD is 1000 wave lengths. Wavelets from all the openings of the grating therefore reach AD in the same phase and therefore conspire to form the bright first order spectrum in the direction EF. But now suppose the direction of AD to be slightly changed so that $BD = 1001$ wave lengths, then CE will equal $500\frac{1}{2}$ wave lengths, and light from B and C will therefore reach AD in *opposite* phases; so also light from the next opening above B will reach AD in opposite phase to that from the next above C, and so on, light from the openings between B and C opposing that from the corresponding openings between C and A. There will therefore be no light of the given wave length in that direction.
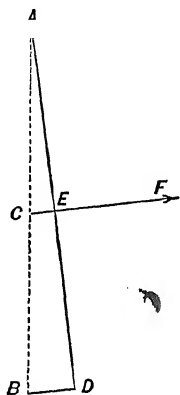
FIG. 595

It thus appears that when BD is 1000 wave lengths there is a bright band of the first order in the direction EF, but this bright band must be exceedingly narrow, for so slight a change in direction of EF as will change BD to 1001 or to 999 wave lengths will take us beyond its limits. *Hence the more lines there are in the grating the narrower will be the bright image due to any one wave length and the closer together two spectrum lines may be and yet be separately distinguishable.*

974. Measurement of Wave Length of Light. Diffraction gratings afford one of the most convenient means of measuring the wave length of light. The grating may be mounted on a *spectrometer*, as shown in figure 596, so that light from the slit S passes through the lens of the collimator and falls upon the grating at G in plane waves. The observer adjusts the telescope T so that the image of some line in the first order spectrum falls on the cross-hairs in the telescope. The telescope may then be moved into the position T' shown by the dotted lines, so that the central bright image (Fig. 589) comes on the cross-hairs; then

the angle between these two positions of the telescope, which is read from the graduated circle, is the angle *dce* or *x* in the small diagram and this is equal to the angle *bac*. But the triangle *acb* is right-angled at *b*, and *bc* is equal to the wave length λ which is to be determined, while *ac* is known from the measurement of the grating and is called the grating space. Representing *ac* by *s* we have $cb = ac \sin x$, or

$$\lambda = s \sin x.$$

From this formula the wave length may be determined when
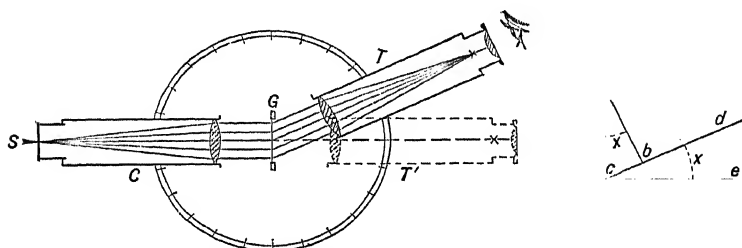
Fig. 596.   Measurement of wave lengths

x has been measured as above described; since *s* is known from the relation $s = \dfrac{1}{n}$, where *n* is the number of lines per millimeter in the grating.

975. Wave Lengths of Some Spectrum Lines.

TABLE OF WAVE LENGTHS

Lines in the Solar Spectrum		Wave Lengths in Millionths of a Millimeter
Extreme limit of visible red.	Fraunhofer's A......	759.4
Deep red.	Fraunhofer's B......	686.7
Red hydrogen line.	Fraunhofer's C......	656.3
Sodium lines.	Fraunhofer's $\Big\{$ D_1.....	589.6
	D_2.....	589.0
Blue hydrogen line.	Fraunhofer's F......	486.1
Nearly limit of visible violet rays.	Fraunhofer's $\Big\{$ H.....	396.9
	K......	393.4

976. Concave Gratings. It was discovered by Rowland that when a grating is ruled on a polished *concave* mirror surface instead of on a flat one *very perfect diffraction spectra may be formed without the intervention of any lenses whatever.* This was a capital discovery, for it was thus made possible to focus the spectra directly on a sensitive plate, and so obtain a photographic map of the lines in the spectrum free from the errors and absorption that lenses introduce.

The mode of mounting a concave grating is shown in figure 597. Two rails AS and SG are fixed at right angles to each
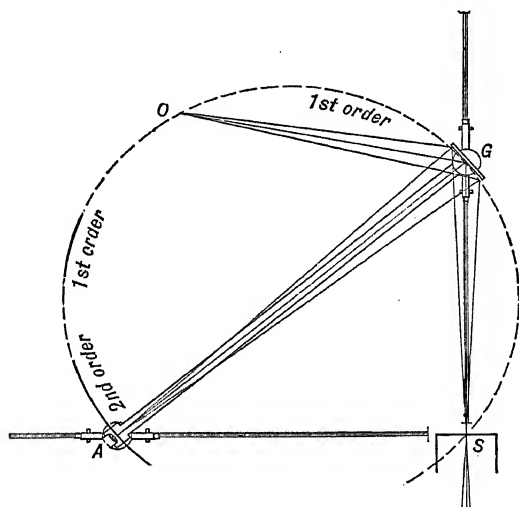
FIG. 597. Concave grating spectroscope

other and a diagonal bar AG, having a length just equal to the radius of curvature of the grating, is mounted on carriages at A and G so that G may be moved toward or away from S along one rail while A moves along the other. The grating is mounted on the diagonal bar at G facing toward the eye-piece or photographic plate holder which is attached to the other end of the bar at A. At S is the slit through which light falls on the grating G. The central bright image of the slit will be found in focus at O, on the circle of which AG is the diameter, and on each side of O the various orders of spectra are formed

in focus on the same circle. In the diagram it will be seen that the instrument is in position for examining the second order spectrum. By sliding A toward S the first order spectrum may be brought in front of the eye-piece.

An important advantage of the spectrum photographs made with this apparatus is that the distances between spectrum lines are proportional to the differences in their wave lengths, so that a scale of equal parts may be made, which when applied to the photograph will give the wave length of every line on the plate.

PROBLEMS

1. Two flat pieces of glass touching at one edge and separated at the other by a thin piece of tinfoil show 30 bright interference bands when examined in sodium light reflected perpendicularly from the thin air film. What is the thickness of the tinfoil?

2. A narrow slit illuminated by light of wave length $600\mu\mu$ gives rise to diffraction bands on a screen 2 meters behind the slit. The two dark bands, one on each side of the central bright band and nearest to it, are just 1 cm. apart. Find the width of the slit.

3. A glass transmission diffraction grating has 50 lines to the millimeter. How far will the first order spectra of sodium light be from the central line when the screen is 6 meters distant?

4. What orders of diffraction spectra will be absent in the spectra produced by a transmission grating in which the bars are exactly equal in width to the spaces between them? See § 972.

DIFFRACTION OF X-RAYS

977. Diffraction of X-rays. The first certain evidence that X-rays (§ 782) are a wave phenomenon in the ether like light but of much shorter wave length was obtained by M. von Laue of Munich in the year 1912. He conceived the idea of using a crystal with its regularly arranged rows of atoms as a diffraction grating. He found that if a narrow beam of X-rays was passed through a crystal of zinc blende with a photographic plate placed a short distance beyond it, on developing the plate a regularly arranged system of spots was obtained as shown in figure 598 which was interpreted as a diffraction pattern produced by very short waves passing through the crystal. Secondary

wavelets arise from each of the regularly spaced atoms within the crystal according to Huygens' principle which on emergence interfere in certain directions and reënforce each other in other directions so that the spotted diffraction pattern is produced. By an analysis of this pattern an idea of the arrangement of the atoms within the crystal can be obtained.

978. X-ray Spectrometer. The English physicists, W. H. and W. L. Bragg, were the first to analyze a beam of X-rays into a spectrum. This was accomplished by " reflecting " the beam from the surface of a crystal.

It happens that crystals such as rock salt have a regular atomic spacing which is about right for X-ray analysis, being in the case of rock salt

FIG. 598. Laue pattern for crystal of nickel sulphate (Bragg)

equal to 2.81$\overset{\circ}{A}$, where $\overset{\circ}{A}$ signifies an *Angstrom*, a unit of length equal to 10^{-8} cms., named after a well-known spectroscopist. The Ångstrom is a very useful unit of measure in crystal analysis
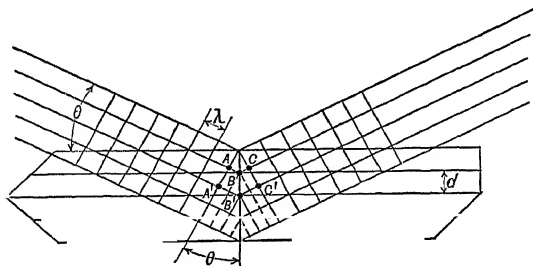
FIG. 599. " Reflection " of X-rays from a crystal

and X-ray work because of its convenient length. The way such a crystal may bring X-rays of definite wave length to focus is shown in figure 599. In this figure the train of parallel waves passes into the crystal and is "reflected" from the different layers. Although spoken of as reflection, it is really diffraction which is produced as · described in § 970. When the waves

emerge, interference between them will result unless the angle θ is such that $AB + BC$ equals one wave length λ and $A'B' + B'C'$ equals two wave lengths, etc. But $AB + BC = 2d \sin \theta$ and $A'B' + B'C' = 4d \sin \theta$. Thus equating either the first to λ or the second to 2λ, the equation for wave length becomes

$$\lambda = 2d \sin \theta. \tag{1}$$

This equation says that for a crystal to reflect X-rays of wave length λ without interference it must be turned until it makes an angle such that $\lambda = 2d \sin \theta$, or $\sin\theta = \dfrac{\lambda}{2d}$. If the X-ray beam contains also radiation of wave length λ', the crystal must be turned to the angle θ' so that $\sin\theta' = \dfrac{\lambda'}{2d}$ and so on for every characteristic wave length the X-ray beam contains. Thus, for every characteristic frequency, the crystal must be at a certain angle to reflect it, and moreover, *the reflected beam* is reflected in a different direction for every characteristic frequency, or X-ray spectrum line. All that is necessary now is a means of detecting these reflected beams of X-rays. This may be done either by their power to effect a photographic plate or to produce ionization (§ 768), that is, by their power to throw electrons off of some of the molecules when passed through a gas. The number of these electrons thrown off per second is easily measured by finding the rate at which they produce a charge on an electrometer.
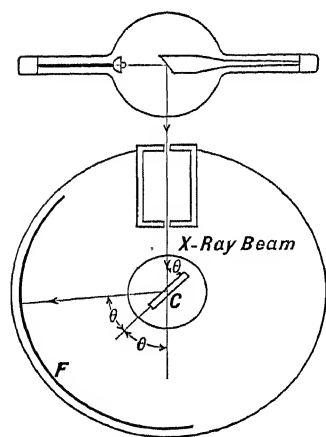
FIG. 600. X-ray spectrometer

A complete X-ray spectrometer is shown diagrammatically in figure 600. The crystal C is mounted so it can be turned through any desired angle. The diffracted X-rays are directed against the film F. The apparatus is inclosed in a vacuum tight chamber if soft X-rays are to be studied, that is, X-rays of comparatively long wave length, because these are absorbed strongly during transmission through a gas. To obtain the

photograph of the entire spectrum on one film, the crystal is slowly revolved until the entire arc of the film is covered. Every time the angle θ passes through a value corresponding to a spectrum line of characteristic wave length, this line is photographed. Special forms of apparatus are used when sharp definition or high resolving power is required.

979. X-ray Analysis of Crystals. W. H. Bragg and W. L. Bragg were the first to use X-ray reflection as a means to determine the exact arrangement of the atoms within certain crystals. A beam of X-rays was analyzed into a spectrum by a crystal X-ray spectrometer as described in the previous section. A beam of monochromatic X-rays of a convenient wave length from this spectrum was isolated from the remaining frequencies by passing it through a slit in a lead screen. This beam of monochromatic X-rays was then "reflected" from the crystal to be analyzed. Although the X-ray beam consisted of only one wave length, as the angle the crystal made with the beam was changed, several positions of the crystal were found at which the X-ray beam was reflected. With one X-ray wave length, equation (1), § 978, shows that only one angle is possible at which reflection can take place for a given spacing of the layers d. The explanation of the several reflection angles found evidently meant that as the crystal was rotated reflection took place from first one and then another set of atomic layers, at different angles from the crystal surface and at different spacings d_1, d_2, etc., like the different systems of rows seen in a corn field while passing by in a train. One such system of atomic layers is shown in the crystal section indicated in figure 599. From a study of the angle of tilt of each such system of layers and their spacing, the atomic arrangement in the crystal could be found. For a complete determination, reflections from more than one face of the crystal are necessary.

The most common atomic arrangement found is a cubical crystal form shown in figure 601, called the face centered cube, which is characteristic of nearly all metals.

In 1916–1917 P. Debye in Europe and A. W. Hull in the United States independently determined crystal structure by passing a beam of monochromatic X-rays through a fine tube of crystal powder of the substance to be analyzed. This method makes use of the fact that in such a powder the crystal fragments are mixed together

Fig. 601.   Face centered cube

with their layers of atoms making every conceivable angle with the X-ray beam so that the several possibler eflections are all obtained simultaneously, each particular reflection taking place from those few crystal fragments which happen to be set just right to satisfy the relation of equation (1), § 978, for that particular layer spacing. For each layer spacing (Fig. 599) the crystal fragments must be set at slightly different angles. Since all angles are present, all possible reflections are obtained. These reflected beams are photographed and appear as a series of lines, an analysis of which may yield the atomic arrangement within the crystals.

Figure 602 shows such a series of lines for the element molybdenum obtained by A. W. Hull by the use of this method.

REFERENCE

A. W. HULL: " The Crystal Structures of the Common Elements," *Journal of the Franklin Institute*, Feb., 1922.

980. Evaluation of the Lattice Spacing of Rock Salt. It will be noticed that after evaluating θ by measurement, equation (1), § 978, still contains both the wave length λ and the crystal lattice spacing d as unknowns. To find λ and thus make possible the evaluation of d for various crystals, a

FIG. 602. X-ray spectrum of molybdenum (Photographed by A. W. Hull)

crystal is selected in which d can be directly calculated. This can be conveniently done for rock salt. Its density is known, its chemical composition is NaCl and since the masses of the individual atoms Na and Cl are known (mass of atom = atomic wt. $\times$ mass of hydrogen atom, see § 628) the number of atoms of Na and of Cl per cubic cm. and thus their spacing can be calculated, since they have a symmetrical cubical spacing. The distance between the horizontal atomic layers of such a crystal is found to be 2.81×10^{-8} cms. or 2.81 Ångstroms. All other crystal dimensions and X-ray wave lengths can be found by comparison, through a knowledge of this value of d.

POLARIZED LIGHT

981. Polarization by Tourmaline. If two plates of tourmaline, cut parallel to the axis of the crystal and of suitable

FIG. 603. Crossed tourmalines

thickness, are placed one upon the other with their axes parallel, light will be transmitted through both plates; but if one is gradually turned on the other the transmitted beam will become fainter until when the two are crossed at right angles there is complete extinction. It is thus seen that light after coming through the first plate of tourmaline is different from ordinary light; for the second plate must have its axis in a particular direction in order to transmit the beam, while in case of ordinary light the transmitted beam is equally intense whatever may be the direction of the crystal axis.

A beam of light having this characteristic is said to be *polarized* and the first plate of tourmaline, which impresses this peculiarity on the light, is called the *polarizer*. The second plate of tourmaline, which reveals the fact that the beam is polarized, is known as the *analyzer*.

982. Direction of Vibrations. If the vibrations in waves of light, like those in sound waves, were perpendicular to the wave front or *in the ray direction*, rotating the tourmaline plate about the ray as an axis would not change its relation to the direction of vibration and consequently the vibrations could not be extinguished in that way, but would pass through both polarizer and analyzer even when they were crossed. We must therefore conclude that *in light waves the vibrations are wholly at right angles to the ray direction.* The French physicist, Fresnel, was the first to draw this conclusion.

983. Nature of Polarized Light. In homogeneous light, or light of one wave length, the vibrations must be in circles or in some form of ellipse or straight line, but in white light they are doubtless very complicated and irregular. But however complicated they may be, each may be conceived as the resultant of two rectilinear vibrations at right angles to each other. For instance, let the path of an ether particle during a short interval be represented by the convoluted line in the diagram, the beam of light being perpen-

FIG. 604

dicular to the paper. It is clear that its actual motion at any instant may be considered as made up of an up and down motion in the direction of the line *ab* combined with a sidewise motion in the direction of *cd*. Any circumstance which would cause one of these component vibrations to be suppressed without affecting the other would leave the particle oscillating along a straight line. It is precisely this which is believed to be effected by the tourmaline plate. Suppose that it absorbs all vibrations at right angles to its axis, while it transmits those which are in the direction of the axis, then the light transmitted through the first tourmaline will all be vibrating in one direction, and if the axis of the second tourmaline is at right angles to the first, no light will get through.
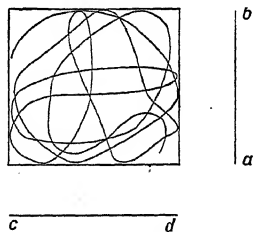
A beam of plane polarized light is therefore believed to be one in which the vibrations all take place in some one direction perpendicular to the ray.

984. Mechanical Illustration. How it is possible for tourmaline to absorb one component of vibration and transmit the other may be seen from the following mechanical illustration. Let a weight of a pound or so be hung as a pendulum from the end of a light strut of wood which reaches out, say, 4 ft. from the wall and is stayed in position by cords *a*, *b*, and *c*, as shown in figure 605. The cords *b* and *c* are somewhat slack and tied into a loop of cord 3 or 4 in. long, which can slip across the end of the strut and is kept in position by three small nails, one above, one below, and one passing through it and limiting the amount of sidewise slip.
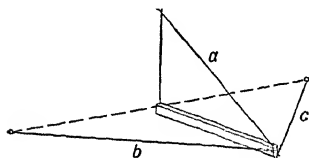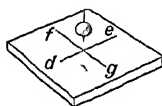
When properly adjusted if the weight is set swinging in the direction *de*, it gives up motion to the strut and soon comes to rest, while if it swings in the direction *fg* the strut is not disturbed and the motion of the pendulum persists. If we now set the weight swinging diagonally or around in a circle, the sidewise component of the vibration is soon suppressed and the weight is left swinging in the direction *fg*.

FIG. 605

Just so, if there is any frictional resistance to the light vibrations in one direction in the tourmaline, the energy of vibrations taking place in that direction will be dissipated in heat and they will be absorbed, while those components of vibration at right angles to that direction may be freely transmitted.

985. Polarization by Reflection and Refraction. When a beam of light falls obliquely on a piece of flat unsilvered glass so that the angle of incidence is about $57\frac{1}{2}°$, the reflected beam is polarized, as may be ascertained by examining the light with a tourmaline plate. It is found that when the axis of the tour-

maline is parallel to the plane of incidence the beam is absorbed by the tourmaline, while if the axis of the tourmaline is perpendicular to the plane of incidence the reflected beam is largely transmitted. And so in general, when light is reflected at the surface of any transparent substance it is found that *for a certain angle of incidence the reflected beam is almost completely polarized. This angle is known as the polarizing angle.*

In this case it is found that *the refracted beam is also polarized in a direction at right angles to that of the reflected beam.* For a crystal of tourmaline having its axis parallel to the plane of incidence absorbs the reflected rays, while it must be held with its axis at right angles to the plane of incidence to absorb most completely the refracted beam.

If the reflected beam contained all those component vibrations of the incident light which are *at right angles* to the plane of incidence while the refracted beam contained all the vibrations *parallel* to the plane of incidence, each beam would be *completely polarized* and they would be equally intense, each having half the energy of the incident beam. But the reflected beam is usually much less intense than the refracted one, and consequently the refracted beam cannot be *completely* polarized, but must contain some of both components of vibration.

In case of glass the beam reflected at the polarizing angle contains only about 9 per cent of the energy of the incident beam. To increase the effect it is common to make use of a pile of thin plates of glass instead of a single reflecting surface. By this device the reflected beam is brighter and the light refracted through the plates is more completely polarized.

Light is polarized in this way also at the surface of opaque substances, such as black glass, which absorb the refracted beam and do not have metallic luster.

Metals and substances having metallic luster reflect both components of vibration and cannot be used to polarize by reflection. Consequently light cannot be polarized by reflection from an ordinary silvered mirror.

986. Brewster's Law. It was discovered by Sir David Brewster that *the polarizing angle for any substance is that angle of incidence at which the reflected and refracted rays are at right*

angles to each other. This is known as Brewster's law of the polarizing angle; it leads at once to the relation

$$\tan p = n$$

where p is the polarizing angle and n is the index of refraction of the substance. For

$$n = \frac{\sin p}{\sin r} \qquad \text{(See figure 606)}.$$

but if the angle between the reflected and refracted rays is $90°$, p and r must be complementary and $\sin r = \cos p$. Therefore

$$n = \frac{\sin p}{\cos p} = \tan p.$$

By the use of this relation the index of refraction of opaque substances, such as dense black glass, may be approximately determined from a measurement of the polarizing angle.

987. Plane of Polarization. Light polarized in the manner that has been described is said to be *plane polarized* to distinguish it from circularly and elliptically polarized light, which will be discussed later.
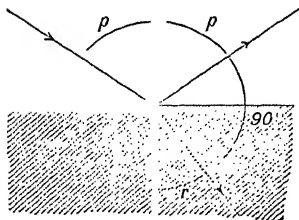
FIG. 606. Polarizing angle

By common consent *a beam of light polarized by reflection* is said to be *polarized in the plane of incidence* or its plane of polarization is said to be parallel to the plane of incidence, while the plane of polarization of the refracted beam is at right angles to the plane of incidence. It is to be understood that *this is simply a convention.*

To find the plane of polarization of any beam of plane polarized light it is only necessary to let it fall on a plate of glass at the polarizing angle and then turn the reflecting plate about the incident beam as an axis until the reflected ray has maximum brightness. The plane of incidence is then the plane of polarization of the incident beam. In this way it may be found

that the plane of polarization of a beam of light transmitted through tourmaline is at right angles to the axis of the tourmaline.

The direction of vibration in plane polarized light is believed to be at right angles to its plane of polarization. This may be inferred from the following case of polarization.

988. Polarization by Fine Particles. When a beam of light shines through a cloud of fine particles it is scattered or diffused to some extent and it is found that the light sent out at right angles to the direction of the original beam is plane polarized. This is easily shown by reflecting a beam of sunlight down into a tall glass jar filled with water made slightly soapy so that it shows a delicate bluish tint. Light is scattered sidewise in all directions so that the path of the beam appears bright, and by means of a tourmaline plate it is found that light coming out in such a direction as *CD* (Fig. 607) is plane polarized, *its plane of polarization being the vertical plane through CD and AB.*

In this case it seems easy to see that on whatever side of *AB* the light may come out, the vibrations in the scattered light must be in a *horizontal direction.* For in the incident beam *AB*, the wave fronts are horizontal and consequently all the vibrations are in

FIG. 607

horizontal planes, and therefore *the vibrations in the scattered light may also be expected to be horizontal*, for the particles which scattered the light are too small to cause an actual turning of the wave front such as takes place in ordinary reflection from an oblique surface.

The above experiment therefore points to the conclusion that the direction of vibration in a plane polarized beam is at right angles to the plane of polarization, as stated in the previous paragraph.

989. Color of the Sky. When the particles are small compared with the wave length of light the shorter waves are most
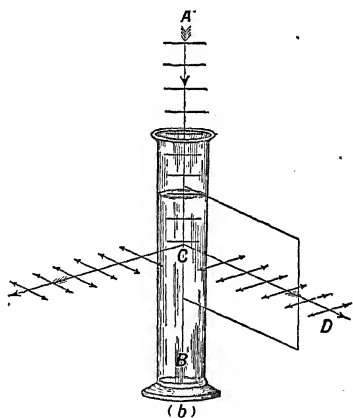
strongly scattered, so that the diffused light is bluish, while the transmitted beam has a larger proportion of the long wave lengths, and therefore appears yellowish or even red.

Thus when we look toward the sun through a thick layer of air filled with fine particles, as at sunset, we see the familiar red and yellow tints; but looking at right angles to the direction of the sun, the diffused light from the sky is bluish and is also found to be polarized.
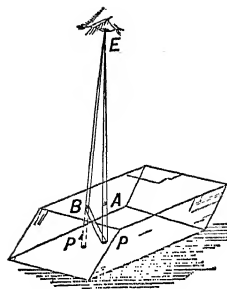
FIG. 608. Double refraction

990. Polarization by Double Refraction. When a crystal of Iceland spar is laid on a printed page, the letters are all seen *double.* A single black dot on the paper appears as two, and if the crystal while lying on the paper is slowly rotated about a vertical axis, one image of the dot is seen to revolve about the other. The paths of the rays of light in this case are shown in figure 608. Light from the black dot at P passes to the eye at E in two beams, one of which PAE is refracted at the surface according to the ordinary law, while the other is bent in an unusual way at B. The first is known as the *ordinary* ray and the other as the *extraordinary*. These two beams are found to be *oppositely polarized*. This may be shown by means of a tourmaline placed on the Iceland spar. If the axis of the tourmaline is in the direction of the line joining the two images P and P' the extraordinary ray PBE is transmitted while PAE is extinguished, while the reverse is true if the axis of the tourmaline is at right angles to the line joining A and B.
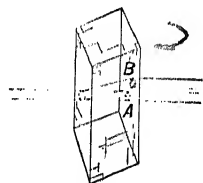
FIG. 609. Double refraction of a narrow pencil of light

When a narrow beam of sunlight falls perpendicularly on one face of a crystal of Iceland spar it is divided into two beams, one of which, the *ordinary* beam, passes straight through, while the other is refracted in an oblique direction in the crystal but emerges parallel to the first at the second face of the crystal, as shown in figure 609. If the incident beam is sufficiently narrow the emergent beams will be separate. otherwise they will overlap. *On*

*testing the two beams with the tourmaline plate they are found to be
oppositely polarized.* The polarization in this case is *complete,*
each beam transmitting only one component of vibration, so that
if the incident light is unpolarized, each of the two beams will
have just one-half the intensity of the original beam.

991. The Double-image Prism of Fresnel. In order to sepa-
rate a beam of ordinary light of the full size of the crystal plate
into two oppositely polarized beams
Fresnel cut the second face of the
crystal obliquely, forming a prism, from
which the two beams emerged in
slightly divergent directions, as shown
in the upper part of figure 610. By
placing a suitable prism of glass in the
reverse position against the prism of
spar, both beams may be bent upward
enough to restore one of them to its
original direction, as shown in the lower
diagram. At a little distance from the
prism the two beams become quite
separate in consequence of their diver-

Fig. 610. Double-image
prism

gence. On looking through such a prism all objects are seen
double. This is one of the best means of obtaining polarized
light where it is desired to transmit both of the two component
beams.

992. Nicol's Prism. When we wish to obtain only one beam
of polarized light, a Nicol's prism may be used. To make such
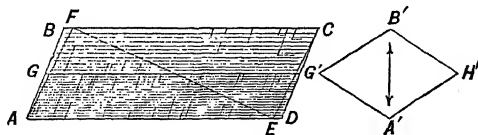
Fig. 611

a prism a long crystal of spar is taken having the form shown in
figure 611, where *ABCD* represents a side view and *A'G'B'H'* an
end view. New end faces *AF* and *EC* are cut, inclined about 3°
more than the natural faces, and the crystal is then divided by an
oblique cut *FE* which is perpendicular to the plane *ABCD* and

also perpendicular to the new end faces. The two surfaces of the cut FE are ground and polished and cemented together with Canada balsam, and the prism is then mounted in a protecting case which permits light to pass through it endwise.

The incident beam I on entering such a prism is doubly refracted, the ordinary ray in the crystal travels with less velocity than in Canada balsam, and meeting the surface FE at an angle greater than the critical angle, is *totally reflected* (§ 878) off to one side, as shown at O. But the extraordinary ray trav-
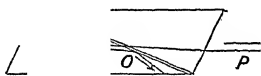
Fig. 612

els in the crystal with a greater velocity than in Canada balsam and therefore cannot be totally reflected and so passes through the prism and emerges as a plane polarized beam in which the direction of vibration is perpendicular to $G'H'$, the longer axis of the rhombus which forms the end of the prism.

A Nicol's prism, or Nicol as it is often called, appears perfectly transparent like clear glass, but the transmitted beam has only half the intensity of the incident one when the latter is not polarized.

993. Double-image Prism as Analyzer. Let a lens L (Fig. 613) be placed in front of an opening at O, through which a beam of
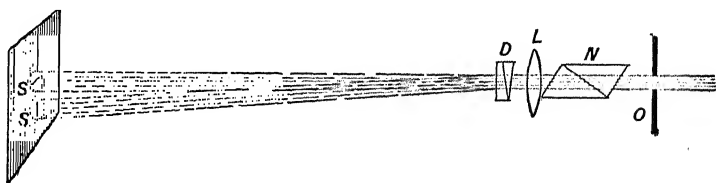
Fig. 613

sunlight passes, so as to form a bright image of the opening on a screen at S. On interposing a double-image prism D two images S and S' are formed. Let us suppose the vibrations in the lower beam to be in the direction of the line SS' joining the centers of the two spots, while in the other they are at right angles to that direction.

Now interpose a Nicol at N so that the light is polarized before

reaching D. Then if the Nicol is in the position shown in the cut, in which it transmits only vibrations in the direction SS', there will be seen on the screen only the spot S'. If the Nicol is slowly rotated about the beam as an axis, the spot S appears, at first faint but growing brighter, while S' grows dimmer, until, when the Nicol has been turned through 90°, S' has vanished and S receives all the light.

To understand these changes let the student in looking at the diagram (Fig. 614) imagine himself looking along the beam of light from O toward the screen, and let N repre-sent the direction of the vibrations transmitted by the Nicol. The double-image prism resolves the vibrations into the two components S and S' which have different velocities in the crystal and are therefore separated, one going to the spot S and the other to S'.

The lines S and S' represent by their lengths the amplitudes of the vibrations in the two beams; and it is evident that as the angle a is increased, the amplitude of S increases and that of S' diminishes until when a is 45°, S and S' will be equal and the two beams of light will be equally bright. Turning the Nicol further causes S to become brighter than S' and when a is 90°, all the light will be transmitted in S, and S' will have vanished.
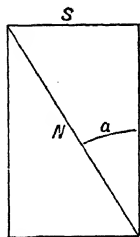
Fig. 614

994. The Wave Surface in Iceland Spar. The Dutch physicist, Huygens, as early as 1690, to explain the double refraction of Iceland spar, advanced the very ingenious idea that a wave of light in such a crystal, instead of spreading out from a center as a spherical wave, divides into two waves, one of which advances as a spherical wave, just as in glass or water, and gives rise to the ordinary ray, while the other wave spreads out as an ellipsoid of revolution and gives rise to the extraordinary ray. He showed that this assumption explained the double refraction of Iceland spar, but he could not explain the polarization of the two beams. This was accomplished by Fresnel, who, in 1821, not only showed that polarization may be explained by the assumption that the vibrations in light waves are *transverse*, but also explained how to account for the separation of the wave in Iceland spar into the spherical and ellipsoidal surfaces conceived by Huygens.

According to Fresnel, the cause of this separation is the fact that *the velocity of light in a crystal depends on the direction of the vibrations in the wave front.* In a crystal there is a certain direction called the optic axis, and in Iceland spar, *waves in which the vibrations are at right angles to the optic axis advance through the crystal with less velocity than waves in which the vibrations are parallel to the axis,* while if the vibrations are neither parallel nor at right angles to the axis the velocity is intermediate.

In the wave surface shown in figure 615 AB is the direction of the optic axis. Vibrations on the surface of the spherical wave sheet are everywhere in the direction of *parallels of latitude* about A and B as poles, and are there-
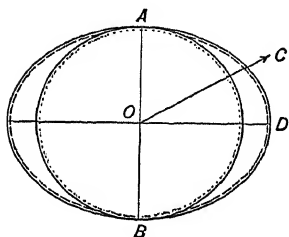
Fig. 615. Huygens' wave surface

fore everywhere at right angles to the direction AB. This wave therefore advances with the same velocity in all directions and must be spherical in form. On the ellipsoidal surface the vibrations are in the direction of *meridians,* consequently at D and at all points on what may be called the equatorial belt of the ellipsoid the vibrations are parallel to AB, at C they are inclined to the axis, while at A they are perpendicular to it, hence the velocity is greatest in directions such as OD, less in the direction OC, and in the direction OA it is the same as for the spherical sheet, for the vibrations in both are perpendicular to the axis.

995. Explanation of Double Refraction. Let a beam of light fall perpendicularly upon the surface of a crystal of Iceland spar, meeting the surface at AB, figure 616, which shows a vertical section through the spar and beam, and let Ax and Bx' be in the direction of the optic axis of the crystal. Then when a wave meets the surface at AB it sets up vibrations at A and B and all intermediate points, which spread out as wavelets in the crystal, each having the form of the wave surface described in the last paragraph. Those components of vibration which are parallel to the line AB go to form the ellipsoidal sheets, while those which are perpendicular to the plane of the diagram form the spherical sheets. The original wave front will thus be separated into two, one of which, CD, is the resultant of the spherical waves and contains vibrations at right angles to the diagram (indicated by dots in the figure), while the other, EF, is the resultant of the ellipsoidal wavelets and has its vibrations parallel to AB (indicated by dashes in the figure). The former is the ordinary ray and the latter the extraordinary. It will be observed that the extraordinary wave EF has greater velocity in the crystal than the ordinary ray, and it moves obliquely, for it must be the envelope of all the ellipsoidal wavelets from A and B and intermediate points.

If the second surface of the crystal is parallel to the wave fronts CD and EF, both beams will emerge perpendicular to the surface (for all points in the wave front EF reach the refracting surface at the same instant) and giving rise to spherical wavelets in the outer medium must advance perpendicular to the surface.

996. Most Crystals Double Refracting. All crystals except those belonging to the so-called regular or cubical system are more or less double refracting. Crystals of the hexagonal and tetrahedral systems have a single optic axis, as in case of Iceland spar and quartz, and are said to be *uniaxial*. While those of the three remaining crystal systems have two optic axes and are called *biaxial*.

997. Double Refraction Produced by Stress. It was discovered by Sir David Brewster in 1816 that when tension or compression is applied to a piece of transparent material such as glass or celluloid, the material becomes doubly refracting, with the optic axis lying in the direction of the applied stress.
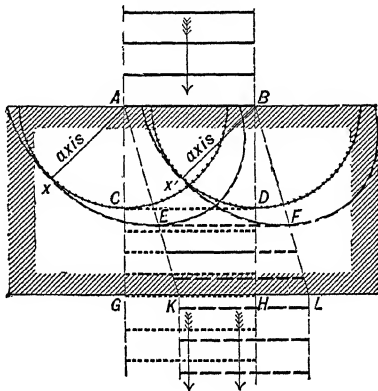
FIG. 616. Double refraction

Since this double refraction effect is proportional to the stress, it gives an optical means of measurement of stresses in transparent materials.

998. Rotation of Plane of Polarization. When a beam of plane polarized light is sent through a crystal of quartz in the direction of its optic axis the plane of polarization is rotated through an angle which depends on the thickness of the quartz and the wave length of the light. Suppose the diagram (Fig. 617) represents the cross section of a crystal of quartz through which light is coming up toward the observer. Let AC be the direction of vibration in the incident beam, then red light may be rotated through the angle AOR and come out vibrating along the direction RR' and violet light being still more strongly rotated may emerge vibrating along VV', with intermediate wave lengths between. If the incident beam is of white light then the emergent beam is also
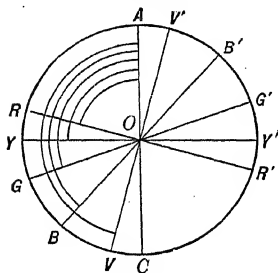
FIG. 617

white, as all the light is transmitted, but if an analyzer such as a Nicol's prism is used, in such a position as to transmit vibrations in the direction RR', the transmitted light will be red, for vibrations at right angles to RR' will be completely cut out, and those in other intermediate directions only partially transmitted. As the analyzer is rotated the tint of the light changes becoming bluish or violet when vibrations in the direction VV' are transmitted while those in the direction RR' are extinguished.

If a double-image prism (§ 991) is used as analyzer the two beams of oppositely polarized light coming from the prism will be of complementary colors when the original incident beam is
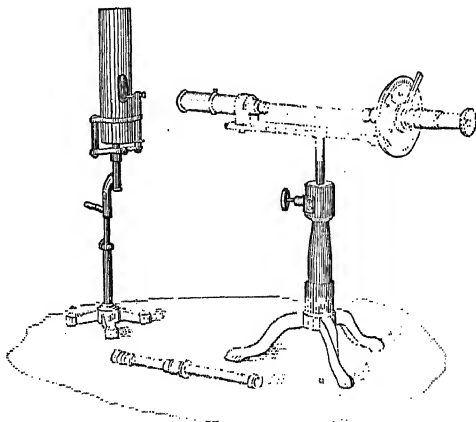
Fig. 618.   Saccharimeter

white, for one will transmit all the component vibrations which are excluded from the other.

Some crystals of quartz rotate the plane of polarization to the right and some to the left, the form of the crystal itself showing to which class a given specimen belongs.

999. Rotation by Liquids. Still more remarkable is the rotation of the plane of polarization by certain liquids, among which may be mentioned turpentine, and solutions in water of tartaric acid, malic acid, and sugar. Both right and left varieties of tartaric and malic acids are known, and a solution containing equal amounts of the two varieties is neutral.

Cane sugar, or sucrose, rotates to the right, but by treatment

with acid it may be broken up into a mixture of dextrose which rotates to the right, and of levulose which rotates to the left. The proportion of cane sugar in a mixture of cane sugar and glucose may be determined by measuring the rotation of the solution both before and after the acid treatment, since the rotation due to glucose is not altered by the process.

An apparatus designed for the exact measurement of the rotation of the plane of polarization by sugar solutions is known as a saccharimeter.

1000. Colors from Crystal Plates in Polarized Light. When polarizer and analyzer are *crossed*, or so placed that the analyzer transmits only vibrations at right angles to those coming from the polarizer, no light will pass through the combination. But if a thin plate of mica or other crystal of suitable thickness is interposed between polarizer and analyzer, as at C in figure 619,
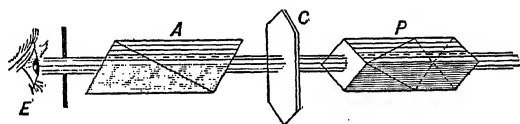
FIG. 619

it may appear vividly colored as seen through the analyzer if the incident light is white. When the crystal plate is slowly rotated, keeping its plane perpendicular to the beam of light, the color is seen to be most intense when the optic axis in the crystal plate makes an angle of 45° with the plane of polarization of the beam, and fades out into darkness when the optic axis of the crystal is either parallel or perpendicular to that plane. The color depends on the thickness of the crystal, and a variety of beautiful colors may often be observed in mica plates in which some parts are thicker than others.

If an ordinary plate of glass is interposed between polarizer and analyzer no effect is observed, the light remains entirely cut off by the analyzer. But if the glass is *in a state of strain*, it acts like a crystal plate and appears bright to the eye at E (§ 997). For instance, if a rod of plate glass is held across the beam at C so that its length makes an angle of 45° with the plane of polarization of the incident beam, on slightly bending the rod, it ap-

pears bright along the edges but dark in the center, for one edge is stretched and the other compressed by the bending, but the center remains unstrained.

So when a piece of glass is heated in a flame and examined between the crossed Nicols, bright regions are seen, due to the strains resulting from unequal heating, but as the heat gradually diffuses through the plate it loses its double refracting power and becomes dark.
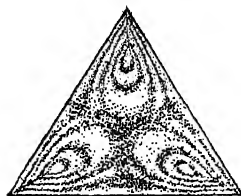
FIG. 620. Strain figure in a triangle of tempered glass by polarized light

Pieces of glass that have been heated and suddenly cooled remain in a strained or tempered state, and when examined with polarized light in the above manner show characteristic patterns as in figure 620.

In this way it may be determined whether the glass for a telescope lens has been thoroughly annealed.

1001. Circular and Elliptical Polarized Light. To understand the production of colors in the case just discussed it will be necessary to consider first what happens when a beam of plane polarized light of *one wave length* passes through a crystal plate. In figure 621 the beam of polarized light is supposed to be coming up toward the eye of the reader. To avoid confusion, the crystal plate and analyzer, instead of being shown superposed on the
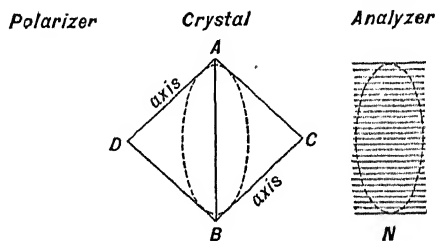
FIG. 621

polarizer as they would actually appear to one looking along the beam, are represented as shifted to one side so that each may be seen separately. The incident light is supposed to be vibrating in the direction shown by the lines at P, with simple harmonic motion, since it is supposed homogeneous. On meeting the crystal it sets up vibrations in the same direction in the face where it enters as represented by the line AB.

But let us suppose that the crystal plate is placed with its axis in the direction DA or BC, at 45° to the direction of vibration in the incident beam. Then the incident vibration, represented in amplitude and direction by AB, may be resolved into the two equal components AC and BC, one of which BC is parallel to the optic axis in the plate while AC is at right angles to the axis. These two components are transmitted with different velocities, and consequently the relation between their *phases* changes as they advance through the crystal.

As the difference in phase of the two components increases the resultant vibration passes successively through the forms shown in figure 622.

When the thickness of the plate is such that one component is retarded one-eighth of a period on the other, the light emerges elliptically polarized, as shown in the second figure in the above diagram. In that case the ana-
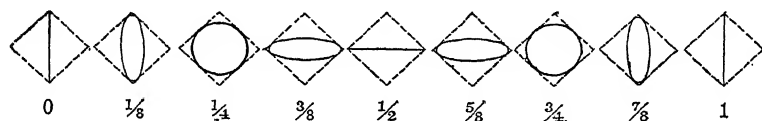
FIG. 622. Resultant forms of vibration when equal simple harmonic components have the difference in phase indicated in fractions of a wave length

lyzer resolves it into vertical and horizontal components and transmits only the horizontal component, as shown in figure 621 at N.

If the crystal plate is of such a thickness that the difference in phase between the two components is a quarter of a complete period, the resultant vibration as it emerges is *circular*. *The emergent beam in this case is circularly polarized, and will be resolved by the analyzer into two components of equal intensity*, one of which will be suppressed and the other transmitted, and there will be no change in the intensity of the transmitted light as the analyzer is rotated.

When the retardation of one component or the other amounts to a half wave length, the emergent light is plane polarized at right angles to the direction of the incident beam and is completely transmitted by the analyzer; while if the relative retardation amounts to a whole wave length the light emerges vibrating just as it entered and is entirely suppressed by the analyzer.

1002. Production of Colors by Polarized Light. When the incident beam of polarized light contains all sorts of wave lengths, as in white light, the crystal plate may appear colored when seen through the analyzer. For, suppose one component of the long waves of red light is retarded a *half wave length* behind the other in traversing the crystal plate, the emergent light will be vibrating at right angles to the incident beam, as shown in figure 622, and will be wholly transmitted by the analyzer.

But in traversing the same crystal plate the shorter waves of violet light may have one component retarded a *whole wave length* behind the other; in

this case the relation of phase is the same in the emergent as in the entering beam, and the beam coming from the crystal is suppressed by the analyzer.

For some intermediate wave length the relative retardation will be $\frac{3}{4}$ of a wave length, and the emergent beam will be circularly polarized and so half transmitted by the analyzer.

On the whole, therefore, the crystal in this case would appear red or orange through the analyzer. But from a crystal plate of twice the thickness both the red and violet waves would emerge vibrating as in the incident beam and would be suppressed by the analyzer, while some intermediate
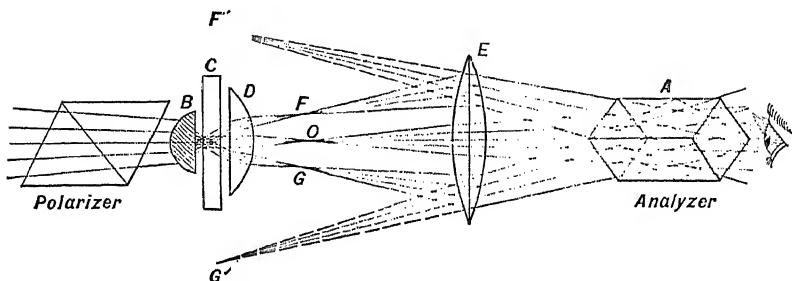
FIG. 623. Optical system of polariscope

wave length would be completely transmitted and the crystal would appear green.

1003. Polarization Figures with Convergent Light. When a thin plate of crystal is examined in an instrument called a polariscope, using a strongly convergent beam of polarized light, a polarization figure is obtained which is
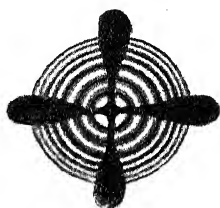
FIG. 624. Polarization figure for uniaxial crystal, perpendicular to axis, Nicols crossed

FIG. 625. Polarization figure of biaxial crystal

of great use to the mineralogist in revealing the optical properties of the crystal.

The optical system of the polariscope is shown in figure 623. The beam of light coming from the polarizer is converged on the crystal by the lens B. On the opposite side of the crystal plate C is a second short-focus lens D, beyond which is the eye lens E and analyzer A.

If the crystal is a plate cut from a uniaxial crystal perpendicular to its axis and if the analyzer and polarizer are crossed, a figure consisting of colored rings intersected by a black cross, as shown in figure 624, is seen at $F'O'G'$ by the observer.

It is clear that rays coming to O in the center of the figure are those that have passed perpendicularly through the crystal section, while rays coming to other points of the figure have passed more or less obliquely through the crystal. Now, the more oblique the rays the greater the thickness of crystal traversed, hence the color seen at any point in the figure depends on the distance of that point from the center at O.

All points on a ring equidistant from O show the same color, because the rays at these points have traversed equal thickness of crystal at an equal inclination to the optic axis.

Rays passing through the crystal in certain directions, however, have their vibrations in such relation to the optic axis of the crystal that they are transmitted without any change in their polarization. All such are cut out by the analyzer and form the black cross.

In figure 625 is shown a more complicated polarization figure produced by a biaxial crystal, such as mica.

REFERENCE ON POLARIZATION

Edwin Edser: *Light for Students.*

Electricity and Light

1004. Magnetic Rotation of Light. When a transparent substance is in a powerful magnetic field a beam of plane polarized light sent through it in the direction of the lines of force has its plane of polarization rotated. This discovery was made by Faraday in 1845 and was the first evidence of a relation between light and electricity and magnetism.

The rotation is usually in the direction of the magnetizing current, or clockwise looking in the direction of the lines of force, though it is opposite in a solution of ferric chloride in water. The amount of the rotation is greatest in substances having a large index of refraction, and in a given substance is proportional to the length of the column and to the strength of the magnetic field.

If the light is reflected back again through the tube the rotation is doubled; that is, the rotation of the plane of polarization produced in this way is the same whether the light passes through the field in the positive direction of the lines of force or the reverse.

This last fact can be explained only by supposing *an actual rotatory motion of some sort* taking place in the magnetic field.

In this respect the magnetic rotation of the plane of polarization is different from that produced by quartz.

1005. The Kerr Effect. Closely related to the rotation discovered by Faraday is the fact, discovered by Kerr, that when a beam of polarized light is reflected from the polished pole of a magnet the plane of polarization is rotated.

1006. Maxwell's Electromagnetic Theory of Light. In the year 1862 Maxwell advanced the theory that light waves are very short electromagnetic waves. Some of the chief arguments for the theory may be thus summarized:

1. The velocity of electromagnetic waves in air is the same as that of light waves. The velocity of a wave depends on the medium in which the disturbance is set up and the kind of disturbance. It is therefore reasonable to suppose that electric waves are the same kind of disturbance as light waves and communicated by the same medium — *the luminiferous ether.*

2. The velocity of light in other media than air has in many cases been found to be equal to the velocity of electric waves in those media.

3. Maxwell showed that electromagnetic waves could not pass through conductors, hence it was to be expected that *conductors would also be opaque to light.* This is strikingly confirmed in the case of metals, for they are the best conductors of electricity and also the most opaque substances known.

4. The electric currents and displacements in the medium transmitting electric waves are parallel to the wave front and at right angles to the direction in which the wave is advancing; and in light waves also the vibrations are parallel to the wave front.

1007. Zeeman Effect. Zeeman of Holland, in 1896, found that when a sodium flame or other luminous gas giving out light of definite wave lengths, as shown by lines in its spectrum, is placed in the powerful magnetic field between the poles of an electromagnet, each single line in its ordinary spectrum is transformed into a group of lines.

1008. Pressure of Light. It was shown by Maxwell in 1873 that if light waves are electromagnetic they must exert a pressure against any surface on which they fall; and that the pressure

against a reflecting surface must be twice as great as against an absorbing one. But the amount of this force is so small that for many years no one succeeded in proving its existence; for in full sunlight, according to Maxwell's theory, the pressure against a reflecting mirror one meter square is only one dyne, or less than the weight of one milligram.

But in 1900 Lebedew in Russia, and in 1901 Nichols and Hull in this country, were able to show that there is such a pressure, and in later experiments to prove that its amount is just what Maxwell's theory indicates.

The pressure of light has been shown by Fitzgerald and Arrhenius to be the probable cause of comets' tails, and Arrhenius has also proposed a very interesting explanation of the Aurora Borealis which depends in part on this same pressure.

RELATIVITY

1009. The Michelson-Morley Experiment. If an observer moves through the air with a body which is giving out sound, he will find that his measurements of the forward and backward velocity of the sound along the path of motion of the body will not be the same. Similarly, it would seem as though motion of the earth through the ether should produce a difference between measurements of the forward and backward velocities of light along the path of motion of the earth through the ether. By means of an interferometer (§ 962) Michelson and Morley compared the velocities of light for different directions at a given point on the earth. No difference in velocity could be detected. It was therefore believed from this experiment that if the earth has any motion through the ether at all the velocity of this motion must be far less than its velocity around the sun *at all times of the year*. This result pointed toward the extraordinary conclusion that the ether, which pervades all space, is stationary relative to the earth or, in other words, that the ether follows the earth in its orbit in preference to all other planets or stars. Since this conclusion is unbelievable, physicists have attempted to account for the difficulty in other ways.

1010. Relativity. Einstein made the assumption that on all planets and stars the Michelson-Morley experiment would lead

to the same result, in other words, that *however fast or in whatever direction an observer moves through space*, the velocity of light as measured by him is always found to be the same. Such an hypothesis is entirely inconsistent with our ordinary conception of wave motion in a medium. How can a wave train moving in a given medium with a definite velocity pass two bodies with the same velocity when one of the bodies is moving faster or more slowly than the other?

Einstein met this difficulty by assuming that there is no such thing as absolute intervals of space and time, but that a length or a time interval is different for one observer from what it is for another, and that this difference between such intervals depends on the motion of the observer. He assumed that this dependence of space and time intervals upon the motion of the observer who measured them was of such a nature as to satisfy the experimental law that the velocity of light should be the same to all observers. This requires that an object moving past an observer A with a velocity v have its dimension in the direction of motion as measured by A *diminished* in the ratio of $\sqrt{1 - \dfrac{v^2}{c^2}}$ to 1, where c = velocity of light. It also requires that a time interval between two events taking place on the moving object as measured by the observer A be *increased* in the same ratio compared with the time interval between the same two events measured by an observer B moving along with the object. He thus set up the revolutionary hypothesis that measured intervals of time and space do not exist as absolute fixed values but that these exist only relative to the observer.

On this hypothesis Einstein then developed the *generalized* theory of relativity, as it is called, an explanation of which is beyond the scope of this book. The most important deduction from this remarkable theory is a new law of gravitation, which is an extremely close approximation to Newton's Law (§ 160). A conclusion from the Einstein Law of Gravitation is that light rays are very slightly deflected as they pass through a gravitational field. This deflection can readily be calculated for a given case and in recent astronomical observations it has been detected and has been found to agree with the calculated value.

RADIATION AND THE QUANTUM THEORY

1011. Introduction. In this chapter a brief introduction will be given to the quantum theory of radiation, a branch of physics which has developed entirely within the twentieth century. The ideas of the quantum theory have assumed such importance in recent years and have presented problems of such extraordinary interest that they have absorbed the attention of many of the ablest physicists of the present day. Modern research in pure physics is largely devoted to the obtaining of data and the unraveling of problems presented by this far reaching and remarkable theory. The science of spectroscopy is founded upon it as is also a large part of X-ray research.

THEORIES OF RADIATION

1012. Radiation. The subject of radiation has already been discussed at some length in the chapter on heat (§ 473–§ 485). See also § 811–§ 824. The process of *emission of energy from a substance in the form of ether waves is called radiation,* and the radiated energy as it exists in the form of ether waves is also spoken of simply as *radiation.* The subject was first introduced under heat because the radiant heat which is felt in front of an open fire, for instance, is a common and well known example of it, and it was first studied in connection with heat phenomena. But it is now known that radiation is a universal phenomenon and that X-rays (§ 782), light waves (§ 846), radiant heat waves (§ 483), or the very long waves of radio (§ 818) are all electromagnetic waves (§ 815) in the so-called ether. The diagram of figure 471 will give some idea of the known range of wave lengths covered by them. The range of wave lengths included in the visible spectrum is seen to be almost infinitesimal compared with the entire known range.

1013. Electromagnetic Theory of Radiation. Before the present century it was felt that very substantial progress had been made in the fundamental understanding of *radiation.* Maxwell in his classic work had shown that ether waves were electromagnetic in nature, being waves of superimposed electric and magnetic fields transmitted with the velocity of light (§ 815 and § 1006). The experiments of Hertz (§ 816) had proved the presence of electric waves of long wave length as required by Maxwell's theory. Lorentz and Abraham and others expanded and developed the theory of electromagnetic waves,

taking account of the part played by the electrons themselves in producing and absorbing radiation. In all of these researches radiant energy was thought of as being propagated in a perfectly continuous flow and absorbed in like manner. The electromagnetic theory was remarkably successful in accurately explaining spectroscopic phenomena, such, for instance, as the simple Zeeman effect (§ 1007) but it fails to explain how electrons maintain their constant speed of revolution in their orbits and at the same time continuously give out energy in the form of waves. How and whence did they derive that continuous supply of energy necessary to keep them revolving at a constant speed?

According to the electromagnetic theory, energy is radiated whenever an electric charge is accelerated. A continuous motion of an electric charge in a straight line, for instance, will produce no radiation. An acceleration or varying motion of the electric charge is necessary. Wireless waves arise from the periodic acceleration of the electric charge as it surges back and forth between the aerial and the earth. There is a definite known source of energy which maintains the waves in this case. The whole phenomenon is accurately explained by the electromagnetic theory. An electron revolving in an orbit should radiate waves also according to this theory since the electric charge is under centripetal acceleration. By Maxwell's theory a train of electromagnetic waves should be produced, of a frequency coincident with the period of revolution. In spite of the difficulty as to how the electron was steadily kept revolving just so, it was, until recently confidently believed by physicists that the periods of the high frequency electromagnetic waves from atoms, which take part in spectroscopic phenomena, were determined by the periods of vibration of electrons in the atoms and that these wave frequencies were therefore coincident with the electron vibration frequencies.

1014. Quantum Theory of Radiation. The basic idea of the quantum theory was first proposed by the German physicist, Max Planck, in a now famous paper in the year 1901. It contained a conception in such complete contradiction to the older electromagnetic theory that it attracted little attention at first. The idea which Planck proposed was nothing less than that energy itself was radiated and absorbed in indivisible grains which he called *quanta*. Planck found it necessary to adopt the idea of quanta in order to explain the way the energy is found to be distributed in the spectrum of a black body (§ 483). His reasons for doing this are more fully explained in § 1039 of this chapter. The idea of quanta has already been briefly discussed at the end of the chapter on heat (§ 484). Energy quanta are primarily radiation phenomena. It is in connection with the absorption and

emission of radiant energy that these minute atoms of energy appear. They are not all of the same size, however, but depend upon the frequency of the radiation. For high frequency radiation such as X-rays, and ultraviolet light, the energy of a quantum is much greater, in

SIZE OF QUANTA

Type of Radiation	Mean Value of ν	$h\nu$ (erg)	$h\nu$ (calorie)
Wireless waves........	10^5	6.554×10^{-22}	1.566×10^{-29}
Heat waves	3×10^{13}	19.66×10^{-14}	4.721×10^{-21}
Visible light	6×10^{14}	39.32×10^{-13}	9.399×10^{-20}
X-rays	3×10^{18}	1.966×10^{-8}	4.721×10^{-15}
Gamma rays.........	3×10^{19}	1.966×10^{-7}	4.721×10^{-14}

fact millions of times greater, than for the longer wave radiation of radiant heat or radio. But even these largest quanta are excessively small particles of energy as is well shown in the accompanying table. For a given frequency of radiation, however, the quanta of energy absorbed or emitted are always of exactly the same size determined by multiplying that frequency by a constant designated by the letter h, called " Planck's quantum of action " as this constant was first proposed and evaluated by Planck in his original paper. It should be clearly understood that h by itself is not a quantum. It is merely that constant which multiplied by a frequency gives the quantum of energy for a given radiation frequency. The size of a quantum is therefore directly proportional to the frequency of the radiation emitted or absorbed as is shown in the table. The proportionality constant h is of course of extraordinarily small size, being equal to 6.56×10^{-27} erg seconds. It is a universal constant entering into all radiation phenomena where quanta appear.

The quantum theory supposes that the absorption and the emission of radiant energy take place in units called quanta, the size of which depends upon the frequency of the radiation.

PHOTOELECTRIC EMISSION OF ELECTRONS

1015. Photoelectric Effect. If strong light which is comparatively rich in ultraviolet rays such as that from an ordinary carbon arc is projected on a plate of zinc which is carefully cleaned and placed in a vacuum and is kept free from coatings of a foreign substance such as an oxygen film, and if this zinc plate is connected to one of the terminals of a quadrant electrometer (**Fig. 311**), the other of which is suitably

grounded (see Fig. 626), it will be observed that the zinc plate becomes positively charged. In other words, the action of the light upon the zinc plate causes it to lose electrons. It has since been discovered that all substances exhibit this so-called *photoelectric emission* of electrons, although there are but few which show it with ordinary visible light, among them the alkali metals, sodium, potassium, lithium, and zinc. This photoelectric effect as it was originally called was first studied quantitatively by the German physicist, Lenard, in work published in 1902.
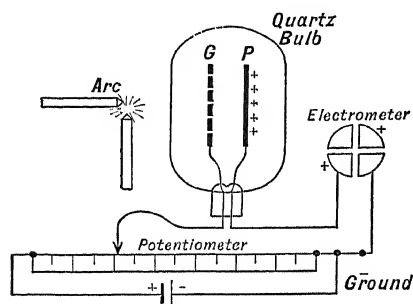
FIG. 626. Photoelectric effect

1016. Laws of Photoelectric Emission. There are two fundamental laws which are found to govern the photoelectric emission of electrons. These are best understood by referring again to figure 626. A grid of wire gauze G (coated with copper oxide to prevent its emitting photoelectrons) is placed in front of the zinc plate P, the grid being so porous that it does not intercept the beam of light projected on P. It is connected with P through a potentiometer (§ 651) so its potential may be raised above or lowered below that of P. By raising the potential of the grid a small amount above that of the plate, all of the *photoelectrons* emitted by the zinc plate are attracted to it so that the total photoelectric current can be measured by noting the rate of accumulation of charge on the quadrant electrometer (§ 574), the latter measuring the positive charge which accumulates on the plate due to its loss of electrons. The photoelectric current is so small that it can be measured easily only by such an electrometer. It is found that *the photoelectric current has a magnitude which is in direct proportion to the intensity of the incident light.* This is the first law of photoelectric emission. Such a law does not suggest anything at variance with the electromagnetic theory of radiation since the effect observed is proportional to the incident energy. But the second law is of a very different character and shows a remarkable departure from what would be expected from the electromagnetic theory. This law concerns the velocity of the emitted electrons and may be stated as follows: *The energy of emission of photoelectrons is determined only by the frequency of the incident light, being directly proportional to this frequency.* The more energetic waves throw out *more* electrons in accordance with the first law of photoelectric emission, but to increase

the velocity of emission of photoelectrons a beam of light of *higher wave frequency* and *shorter wave length* must be used, even though it may be a much feebler beam than the former.

Although comparatively simple to state, this second law of photoelectron emission is only arrived at experimentally by taking extraordinary precautions besides taking account of the characteristic attraction of the plate itself for electrons in interpreting the result.

1017. Velocity of Emission of Photoelectrons. The velocity of emission of photoelectrons is found by observing the negative potential necessary to prevent electrons emitted from the plate from reaching the grid. With a fixed intensity of monochromatic light incident on the plate, that is, light of only one wave length from a single spectrum line, it is found that as the potential of the grid is gradually

FIG. 627. Curve of photoelectric emission showing cut-off potential

decreased below that of the plate, fewer and fewer electrons reach the gauze grid until a definite negative potential is reached where it receives no electrons at all. · This is illustrated by the curve of figure 627 which shows a cut-off potential of V_m volts. This cut-off potential is evidently a measure of the energy of the fastest photoelectrons.

The repulsion of electrons by the grid at this potential is just sufficient to prevent the fastest moving ones from reaching it. They have therefore just lost their kinetic energy of emission which equals $\frac{1}{2}mv^2$ (§ 107) where m is the mass of an electron and v the velocity with which it was emitted. This lost kinetic energy must be exactly equal to the work done by the electron as it moves against the negative field. This work is equal to eV where e is the charge of the electron and V the applied negative potential. This follows from the definition of potential according to which a difference of potential V between two points is measured by the work done upon unit charge as it is moved from one point to the other (§ 571). Therefore the work done upon a charge of e units is equal to eV in moving it from one point to the other. Equating this to the corresponding kinetic energy lost,

$$eV = \tfrac{1}{2}mv^2. \tag{1}$$

Since e and m are known, and V can be measured, the velocity v of the photoelectron can be calculated from this equation. This equation gives the velocity of motion of the fastest moving photoelectrons only.

In the photoelectric effect it is this *maximum* emission velocity which is significant as will be seen from the discussion which follows.

1018. Einstein's Photoelectric Equation. It was first proposed by Einstein in the year 1905 that the photoelectric emission of electrons might be explained on the basis of the quantum theory. According to his hypothesis each individual electron which was emitted received just one quantum of energy from the incident radiation of magnitude $h\nu$. Since every metallic surface has a characteristic attraction for electrons a small amount of work w must be done to pull an electron away from the surface against this attraction (§ 778). Therefore the quantum of energy imparted to the electron is used up partly in overcoming this attraction so that the kinetic energy of emission $\frac{1}{2}mv_e^2$ is that which remains after the electron leaves the surface of the plate during photoelectric emission. Expressing these facts in mathematical form, there results the equation:

$$\tfrac{1}{2}mv_e^2 = h\nu - w. \tag{2}$$

In words, this equation says that the energy of emission of an electron from a plate due to the action of radiation upon it equals the energy of one quantum of the incident radiation minus w, the work necessary to pull the electron out of the plate. It is evidently assumed that each individual quantum delivers its entire energy to an electron during this process. Just how this is done remains as yet one of the puzzles of modern physics.

Using equation (1), equation (2) can be put in the form:

$$Vc = \tfrac{1}{2}mv_e^2 = h\nu - w. \tag{3}$$

When this equation was first proposed by Einstein, exact experimental evidence for it was wanting and some years passed before it was generally believed, so radically new was the conception which it contained. Nevertheless, during the few years that followed, it has been tested by the most careful experiments, notably by those of Millikan (§ 1019), and when every conceivable source of error is eliminated this equation is found to express the observed facts, almost exactly. It is, in fact, a mathematical statement of the second law of photoelectric emission introduced in § 1016.

This equation, simple though it is, is unquestionably one of the most important in modern physics.

1019. Verification of Einstein's Equation. In the verification of Einstein's equation a method for measuring the velocity of photoelectrons similar to that described in § 1017 was used where the *maximum* velocities only are evaluated. At first thought the question

arises, why are large numbers of slower moving photoelectrons found to be present when they all receive equal increments of energy $h\nu$. This is probably because the photoelectrons, besides being emitted in all directions, suffer considerable interference from neighboring atoms and molecules. The *maximum* velocities observed represent the few photoelectrons emitted normal to the surface of the plate with no interference, and therefore it is the fastest observed photoelectrons which give a true measure of the energy of the quanta.

Figure 628 is a curve of photoelectron emission from a plate of sodium, as obtained by Millikan, which shows how the voltage necessary to stop the fastest moving photoelectrons increases as the frequency of the incident
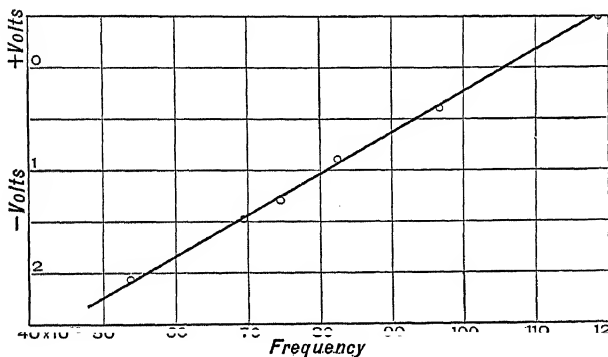
FIG. 628. Photoelectric emission from sodium (Millikan)

light used is increased. The ordinate can be expressed as the velocity corresponding to the voltage V as found from equation (1) if desired. It is expressed in volts merely for convenience. Millikan found that h as found from such curves as these agreed closely with the value originally assigned to it by Planck, which is indeed a convincing experimental verification of Einstein's equation. The work w, necessary to remove an electron from the metal surface, is a constant quantity, for all frequencies used (§ 778). It merely displaces the curve downwards a certain amount and drops out of account in the evaluation of h.

1020. Photoelectric Action of X-rays. As implied by Einstein's equation photoelectric action is not confined to the relatively low frequencies such as used by Millikan in his experiments. Direct evidence of this is obtained when a piece of metal is exposed to X-rays. Photoelectrons of high velocity are emitted from the metal with a maximum velocity which may be calculated from the maximum X-ray wave frequency. Photoelectrons produced by X-rays have far more kinetic

energy, however, than those obtained in Millikan's experiments, in fact more than a thousand times as much, as is seen from Einstein's equation, since the frequency of X-ray radiations is more than 1000 times that of visible light (see Fig. 472).

It is to be observed that for X-rays, the quanta $h\nu$ are of such a magnitude that w, the work required to remove an electron against the surface attraction of the metal in equation (3), is so small, comparatively, that it is entirely negligible in experimental work, so that Einstein's equation takes the simple form

$$\tfrac{1}{2}mv_e{}^2 = h\nu. \tag{4}$$

1021. Inverse Photoelectric Effect. Important and striking as the "direct" photoelectric effect is, its inverse is even more far-reaching in furnishing experimental data for quantum investigations. In fact the whole subject of modern spectroscopy is founded upon it. The term "*inverse photoelectric effect*" is not in general use but as the name implies, it refers to *radiation emitted by matter bombarded by electrons*, being the converse of the former. A well known instance is the production of X-rays by cathode ray bombardment in an X-ray tube. Furthermore, when the inverse photoelectric effect is quantitatively examined, Einstein's equation is again found to be satisfied. For example, where electrons strike an X-ray target with a known kinetic energy as calculated by equation (1) where V is the tube voltage, it is found that the radiation emitted, however complex in character it may be, has an upper frequency limit ν_e such that the quantum $h\nu_e$ carrying this radiation is just equal to the electron kinetic energy, or, as before,

$$h\nu_e = \tfrac{1}{2}mv_e{}^2.$$

Here again Einstein's equation takes the simple form in which w is neglected because of the relatively large size of the quanta produced. Thus in an X-ray tube the higher the voltage applied to the tube the higher the X-ray frequencies obtained. The reversal of photoelectric effects produced by low frequency visible rays, such as studied by Millikan, has never been observed, probably because the radiation produced is of such small intensity as to be unmeasurable.

1022. Experiments of Hunt and Duane. A very interesting check on Einstein's equation for X-rays was obtained from experiments of Hunt and Duane, the results of which are shown in the curves of figure 629. Each curve was obtained by studying the effect of variation of voltage across the X-ray tube upon the intensity of a particular frequency in the X-ray spectrum. An X-ray spectrometer crystal

(§ 978) was adjusted to "reflect" the desired frequency. This reflected beam was passed through an *ionization chamber* by which its power to ionize air (§ 769), and thus its intensity, was recorded in the arbitrary units shown. Each curve required a particular adjustment of the spectroscope, and is marked with the X-ray wave length obtained.

The important information which these curves give is that every wave length has a definite "threshold value" of excitation voltage, that is, voltage through which the exciting electrons drop against the
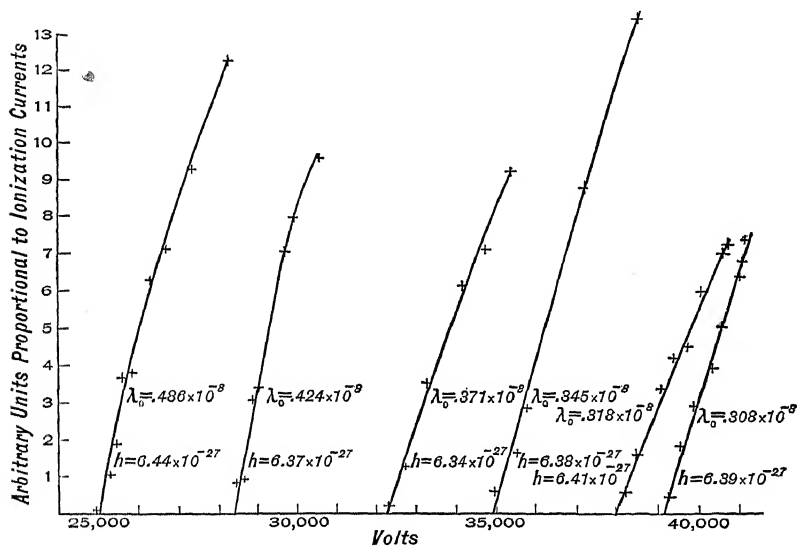
FIG. 629. Hunt and Duane's curves

target material, below which no radiation of this wave length can be obtained. Furthermore, this voltage is in each case almost exactly that required by Einstein's equation.

This can be verified easily by substituting in Einstein's equation in the form given neglecting w,

$$Ve = 300h\nu \qquad (5)$$

where V is measured in volts (hence the factor 300, see § 741) $e = 4.774 \times 10^{-10}$ electrostatic units and the frequency $\nu = \dfrac{c}{\lambda}$ where c the velocity of light $= 3 \times 10^{10}$ cms. per sec. and λ = wave length in cms.

For instance, in the case of the left hand curve where $\lambda = .488 \times 10^{-8}$ cm., $V = 25,000$ volts at the point where the radiation intensity vanishes. Then substituting in equation (5) and solving for h, its value is found to be 6.44 $\times 10^{-27}$ C. G. S. units, which is very close to the value 6.56×10^{-27} obtained by Millikan.

Thus it is seen that the emission of X-ray frequencies obeys this fundamental quantum equation. This agrees with the statement at the end of § 1014 according to which radiation is emitted as well as absorbed in quanta. Furthermore this emission and absorption obeys Einstein's equation in the cases described.

THE BOHR ATOM

1023. The Spectral Series of Atomic Hydrogen. It has already been explained that a gas such as hydrogen has a spectrum consisting of series of lines which are brilliantly colored within the visible range.
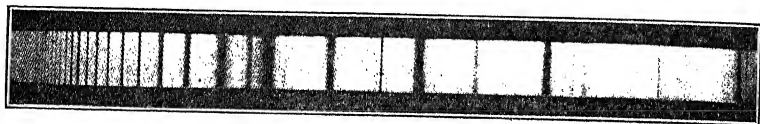
FIG. 630. Balmer series of hydrogen (absorption spectra of the star Zeta Tauri — *Foote & Mohler*)

A remarkable numerical relation was discovered by Balmer (1885), and amplified by Rydberg (1890) and Ritz (1908), from which the exact frequencies of whole series of spectrum lines of this type could be calculated. For the case of hydrogen, this relation takes the very simple form:

$$\nu = Rc \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{6}$$

where ν is the spectrum line frequency, c is the velocity of light and R a constant called the Rydberg constant and where the different frequencies are found by giving n_1 and n_2 a series of values of whole numbers. For instance the well known Balmer series of spectrum lines of atomic hydrogen, so named after its discoverer, is obtained by giving to n_1 the value 2, and to n_2 the series of values 3, 4, 5, etc., up to infinity. A photograph of this series is shown in figure 630. This is an absorption spectrum (§ 936) of the star Zeta Tauri. It is seen that these lines (that is, their frequencies) become closer and closer together, as the series approaches its limit corresponding to the case where $n_2 =$ in-

finity. This is called the series limit. Every spectrum series has a limit of this kind. The first few lines of such a series are usually comparatively easy to observe, but the lines soon become close together and fainter as the limit is approached, although the limit itself is often quite clear. Other series in the spectrum of this atom besides the Balmer series are to be expected on the basis of equation (6). There should be a series for which $n_1 = 1$ and n_2 takes the values 2, 3, 4, etc. This series has been identified principally by Lyman, and is called the Lyman series. Careful and difficult experimental work was necessary to obtain the lines of this series because these lines fall far in the ultraviolet region. Radiation in the frequency range of this series (from $\lambda = 915$ Å to $\lambda = 1220$ Å) is powerfully absorbed by gases of *all* kinds, so the spectrometer and apparatus for their measurement must be isolated in a vacuum chamber.

On the other hand, two other series of longer wave lengths than the Balmer series have been located in the infra red region. The best known of these is called the Paschen series after its discoverer. This series is obtained from the frequency equation by setting n_1 equal to 3, and assigning to n_2 the values 4, 5, 6, etc. The strongest lines of the Balmer series fall in the visible region, so this series was identified before the Lyman and Paschen series.

1024. Bohr's Frequency Condition for the Hydrogen Atom. The significance of this relation and similar relations for atoms other than hydrogen was for many years an outstanding puzzle for spectroscopists, as it defied explanation on the basis of the electromagnetic theory, whatever structure was assumed for the atom. About the year 1913, the Danish physicist, Niels Bohr, following Planck (§ 1014) abandoned the electromagnetic theory and took the radical step of invoking the quantum theory as the basis of explanation of spectrum line frequencies.

In the case of the hydrogen atom, he took the spectrum frequency formula of equation (6) and multiplied both sides by Planck's constant h, which made this equation take the form

$$h\nu = \frac{Rhc}{n_1^2} - \frac{Rhc}{2}. \tag{7}$$

The left hand member of this equation, $h\nu$, becomes a quantum of energy, so that instead of representing a series of spectrum line frequencies, as before, this equation gives a series of quanta, one corresponding to each spectrum line. For the Balmer series of hydrogen, for instance, when n_1 is given the value 2 and n_2 the series of values 3, 4, 5, etc., equation (7) gives a series of values of the quantum $h\nu$, one

corresponding to each line of the Balmer series. Since $h\nu$ is a certain amount of energy, the two terms on the right hand side must each represent energy. According to Bohr, each of these terms represents definite energy states of the atom itself and when the atom changes from a state of energy represented by the second term to a state represented by the first term on the right hand side of this equation, the energy thus lost by the atom *goes out* in radiation as one quantum of size $h\nu$. If the energy of the atom changes from the first to the second state, its energy increases by one quantum $h\nu$, which is therefore *absorbed* from incident radiation from outside.

1025. The Bohr Atom.   Bohr then undertook to obtain a physical interpretation of equation (7). Such an interpretation requires a definite type of atom whose energy is always equal to some one of a series of definite values. These energy values or energy levels must be such that for the case of hydrogen the energy shifts between these levels give series of energy changes identical to those given by equation (7).

He assumed the Rutherford type of atom (§ 797) with a relatively massive positively charged nucleus with electrons revolving about it in orbits, the number of these electrons being equal to the atomic number.

He assumed Coulomb's law of attraction to hold between the nucleus and electrons and centrifugal force to hold the electrons in their orbits against the nuclear attraction.

The energy of such an atom is determined by the energy of the electrons in their orbits, being greater the larger the orbits.

He found for the case of the hydrogen atom with one revolving electron that if it was assumed that the angular momentum of this electron was equal to $\dfrac{nh}{2\pi}$ (Appendix III) the series of orbits determined by the series of integral values of n gave exactly the series of energy values required by equation (7). It is surprising that so simple a relation as this latter should give the correct series for hydrogen. The Bohr atom thus immediately attracted wide attention, and was found to be capable of extension to atoms of greater complexity than hydrogen.

It can be shown that it necessarily follows from these assumptions that the Rydberg constant R becomes

$$R = \frac{2\pi^2 me^4}{ch^3}. \tag{8}$$

For an elementary presentation of the mathematics of the Bohr atom, see Appendix III. The Bohr atom violates the electromagnetic theory in two ways. It postulates absorption and emission of radiation in

quanta instead of in a continuous flow. It presupposes that electrons may revolve in their atomic orbits without producing radiation. On the basis of the electromagnetic theory the revolving electrons should radiate energy and move in toward the nucleus in a spiral as the energy decreases. Bohr therefore takes the position that the electromagnetic theory does not apply to atomic radiation and absorption processes.

1026. Remarks about the Bohr Atom. The Bohr atom for hydrogen is thus seen to consist of a nucleus, with a positive charge equal but opposite to that of an electron, with one electron revolving about it. The lowest possible energy state according to the above formula (7) is the one for which $n = 1$. No energy state smaller than this can occur according to Bohr's theory, since n cannot have an integral value smaller than unity. If this were so, one of the terms of equation (7) would become infinite, giving infinitely large quanta. When the electron is in this smallest orbit where $n = 1$ the hydrogen atom is considered to be in its normal state. When the electron is lifted out of this orbit by absorption of a quantum of energy to some orbit of greater energy, to that for which $n = 4$ for instance, the electron strives to return to its normal orbit. It remains in the orbit $n = 4$ for an instant and then drops perhaps to $n = 2$, remains there for an instant and at length drops back again to $n = 1$ again, unless lifted up to a higher orbit by some new quantum of radiation before it can return.

It is clear from this consideration of how quanta are absorbed and emitted, that the only way all lines of a spectral series can appear simultaneously is that many atoms are undergoing this absorption and emission of quanta of various sizes simultaneously so that the total effect is the appearance of all possible lines at once or complete spectral series.
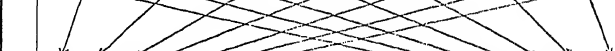
There are other ways in which the Bohr atom conforms to experimental facts besides its ability to predict the spectral series of hydrogen as described.

If the distance a of the electron from the nucleus in the normal state of the atom when $n = 1$ is calculated by equation (11), Appendix III, it comes out 1.1×10^{-8} cm. which is about the value the radius of the hydrogen atoms should have from predictions of the kinetic theory.

1027. Unoccupied Orbits. All of the possible orbits of the hydrogen atom in the previous discussion except that of the undisturbed electron are occupied only temporarily by the electron during the absorption and emission of radiation. These orbits are spoken of as *unoccupied* or *virtual* orbits. The unoccupied orbits are infinite in number, corresponding to the infinite number of possible values which may be assigned to n. The permanent orbits are of course equal in

number to the electrons revolving about the nucleus, for any atom, and are thus equal in number to the atomic number of the atom. Hydrogen has one permanent orbit, and uranium 92 of them. This important distinction between permanent and unoccupied orbits should be clearly kept in mind in thinking of the Bohr atom.

1028. Spectra of the Hydrogen Type from Other Atoms. It is seen from the accompanying periodic table that helium is of atomic

O	I	II	III	IV	V	VI	VII
He 2	Li 3	Be 4	B 5	C 6	N 7	O 8	F 9
Ne 10	Na 11	Mg 12	Al 13	Si 14	P 15	S 16	Cl 17

O	Ia	IIa	IIIa	IVa	Va	VIa	VIIa	VIII			Ib	IIb	IIIb	IVb	Vb	VIb	VIIb
Ar 18	K 19	Ca 20	Sc 21	Ti 22	V 23	Cr 24	Mn 25	Fe 26	Co 27	Ni 28	Cu 29	Zn 30	Ga 31	Ge 32	As 33	Se 34	Br 35
Kr 36	Rb 37	Sr 38	Y 39	Zr 40	Nb 41	Mo 42	43	Ru 44	Rh 45	Pd 46	Ag 47	Cd 48	In 49	Sn 50	Sb 51	Te 52	I 53
X 54	Cs 55	Ba 56	La-Lu 57-71	Hf 72	Ta 73	W 74	Re 75	Os 76	Ir 77	Pt 78	Au 79	Hg 80	Tl 81	Pb 82	Bi 83	Po 84	85
Rn 86	87	Ra 88	Ac 89	Th 90	Pa 91	U 92											

Periodic table showing atomic numbers (hydrogen = 1)

number 2, lithium is of atomic number 3, etc., up to uranium which is of atomic number 92. It will be recalled that the atomic number is the same as the number of elementary positive charges carried by the nucleus, that of hydrogen being simply 1, that of helium 2, that of lithium 3, etc. If in the preceding equations of the Bohr atom for hydrogen the nuclear charge E is replaced by $2E$, as shown in Appendix III, equation (7) becomes

$$h\nu = 4Rch\left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right). \qquad (9)$$

This equation, however, does not give the energy states of the neutral helium atom, because the latter has *two* electrons revolving about its nucleus. Equation (7) was developed for the case of a *single* electron revolving about a nucleus. Thus it is reasonable to expect that equation (9) will give the energy states for a helium atom with only one electron revolving about its nucleus, that is, for the ionized (§ 769) helium atom. This has actually been proved to be the case by experiment, the lines of ionized helium being first identified by the English physicist, Fowler. Accurate measurement shows, however, that

the Rydberg constant of equation (9) as applied to helium differs
slightly from that of equation (7) as applied to hydrogen. It is re-
markable that this discrepancy is exactly explained by the more exact
forms of equation (7) and (9) which take account of the slight motion
of the relatively massive nucleus as an electron revolves about it which
was neglected in the derivation of these equations as presented in
Appendix III. The spectrum of *neutral helium* with its two revolving
electrons is very much more complex than that of ionized helium or
hydrogen. When $E = 3e$, equation (7) becomes

$$h\nu = 9Rch\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{10}$$

Following the previous reasoning this equation should give the energy
states for a lithium atom with 2 of its 3 electrons removed, that is, for
doubly ionized lithium. These lines have never been identified be-
cause they lie far in the ultraviolet where measurement is very dif-
ficult.

1029. The General Bohr Theory. The presentation of the Bohr
theory which has been given assumes circular orbits only, because of
the comparative simplicity of the treatment required. In his more
complete theory Bohr assumes the orbits to be of elliptical form, of
various degrees of eccentricity, that is, ellipses of varying lengths com-
pared with their breadths. Orbits of certain definite eccentricities
only are permitted, each eccentricity corresponding to a definite
energy level. Furthermore the ellipses may be in different planes.
These details of the more complete Bohr theory, although of interest
to the spectroscopist in furnishing a more complete picture of the
atom, must not be regarded as final. Serious difficulties still remain,
the study of which is beyond the scope of this book. The great con-
tribution of the Bohr theory is not so much in the systems of electron
orbits proposed as in bringing to clearness and reality the existence of
series of definite energy levels of the atom, and in showing how absorp-
tion of radiation by an atom is attended by a change from a lower to a
higher energy level of that atom, and emission of radiation is attended
by a change from a higher to a lower energy level of the atom, and that
to each such transformation a definite quantum of radiant energy is
absorbed or emitted.

From this point on, orbits will not be referred to, but the term
energy level will be used instead. It is the energy levels of the atom
which are fundamental. In experiments on the absorption and emis-
sion of radiation it is these energy levels which are actually observed.
They thus have a reality which must always remain whatever be the

mechanism within the atom which produces them.   The series of energy levels for the hydrogen atom is shown in figure 631, the arrows showing the energy shifts which produce the Balmer series as calculated from equation (7).
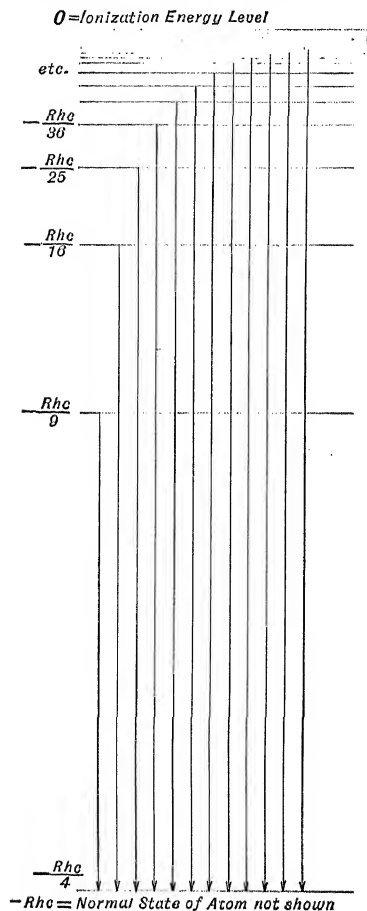
FIG. 631.   Hydrogen energy level series

0 =*Ionization Energy Level*

etc.

$-\dfrac{Rh c}{36}$

$-\dfrac{Rh c}{25}$

$-\dfrac{Rh c}{16}$

$-\dfrac{Rh c}{9}$

$-\dfrac{Rh c}{4}$

$-Rh c =$ *Normal State of Atom not shown*

SIMPLE CASES OF ATOMIC SPECTRA

1030. The Inert Gases.   It is well known that the neutral atoms of the inert gases, helium, neon, argon, krypton, etc., are extremely stable.   This is inferred partly from their chemical inactivity.   They refuse to combine with other atoms to form compounds, showing that the outside electrons are held with great stability, leaving no stray fields of force, to speak of, whereby they might attach themselves to other atoms.   They are thus monatomic, not even clinging together in pairs to form diatomic molecules as most gases do.   They all exist as gases, until extremely low temperatures are reached when they finally liquefy or solidify.   This indeed does show the presence of very slight stray fields of force, but fields so weak that the atoms cling together to form liquids only when the agitation due to temperature has largely disappeared.   For this reason they occupy their peculiar positions in the periodic table. The spectra of these elements are very complicated, because of the large number of electrons in the outer shell of each, all of which are involved to some extent even for the most easily excited spectra, such as produced by an ordinary electric arc discharge through these gases.

1031. The Spectrum of Sodium.   Such is not the case for sodium, however.   The spectrum of this element is one of the simplest of any of

the elements. The reason for this is that its atomic number is just one above that of the very stable element neon, mentioned in the previous section, its one extra electron remaining outside of the other ten which keep the stable neon formation, unless violently excited by an electric spark. The spectrum produced by an arc discharge through sodium vapor (or the arc spectrum as it is called) involves the single outside electron only, and hence its comparative simplicity. The various energy levels corresponding to the lifting of this electron to its various possible orbits as determined by Bohr's quantum condition are shown
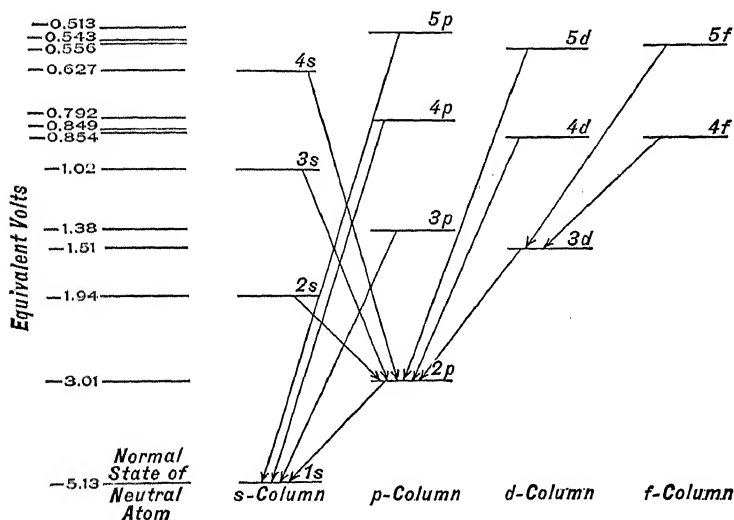
Fig. 632. Energy levels of the sodium atom

in figure 632. The spectral series produced arise from the energy changes from one state of the atom to another shown by the arrows. It is a curious fact, not hitherto mentioned, that not all possible changes (from a higher to a lower energy level) actually occur, such as a change from column of energy states s to column d. Shifts are seen to take place between adjacent columns only. The reasons for the separation of the energy levels of the sodium arc spectrum into the 4 adjacent columns only, are very well discussed in more advanced works on this subject. See in particular the references at the end of this chapter by K. K. Darrow and A. Sommerfeld.

Simple as it is, the series of energy levels of sodium as given in figure 632 is seen to be more complex than that of hydrogen given in

figure 631. Strictly, it is believed that the energy levels of hydrogen also should be separated into columns like those of sodium, and that energy changes between adjacent columns only should be permitted. For hydrogen, however, according to the Bohr theory, the adjacent columns would have energy levels almost identical with one another, so that the spectral series (arising from transformations between adjacent columns) would be the same between any two columns. Therefore a single series as shown in figure 631 with possible energy changes from one level to another in this figure is sufficient to give all of the spectrum lines actually observed, besides giving the possible energy changes given by equation (7). Equations similar to (7) exist for the sodium series also, one for each series of changes between the adjacent columns of figure 632.

The series corresponding to transformations from column s to p shown by the arrows is called the *sharp* series, that from p to s, the *principal* series, that from column d to p the *diffuse* series, that from f to d the *Bergmann* series or the fundamental series. These series were named by spectroscopists many years before the Bohr theory was developed with its idea of energy states and emission and absorption of radiation in quanta. The principal series was so called because of its greater intensity and prominence compared with the others in the sodium group of atoms. According to the Bohr theory this greater prominence is to be expected since the state marked $1s$ in figure 632 is the state corresponding to the normal position of the outer electron which produces these series, to which the electron is always striving to return. All of the other states correspond to levels temporarily occupied by the electron when lifted temporarily out of its normal level by the excitation of an arc discharge, for example. The principal series arises from transformations from the second or p column down to this most probable normal $1s$ level. It is thus natural to suppose that the principal series transformations would be the most probable. The lines of the other series are much less intense. Still other series are possible and are observed, although comparatively faint, such as a p-d series or a d-f series corresponding to changes from column p to d and d to f. Still other energy levels exist for sodium, but they are not so well known and do not appear in the figure.

1032. Arc and Spark Spectra. If vapor of the element magnesium is placed between the electrodes of an ordinary arc this element is found to produce a complex system of spectrum lines which cannot be resolved into systematic series which are as simple as those of sodium shown in figure 632. This is believed to be due to the fact that two electrons instead of one are taking part in the production of the system of lines, since the atomic number of magnesium is 12, or one higher

than sodium, and it has two loosely bound electrons outside of the inner 10 electrons which make up the stable neon arrangement. If, however, magnesium is strongly excited between the electrodes of a spark discharge of considerably higher voltage than that of an arc, the spectrum changes, and becomes very similar to that of sodium. This is due to the complete removal of the 12th electron, so that only one loosely held electron remains outside of the neon electron system, giving an atomic structure like that of sodium. The spectrum of *singly ionized* magnesium, as this is called, is thus very similar to that of sodium, except its series are of four times the frequency owing to the nucleus having an extra unit of positive charge on it. Again, if aluminum of atomic number 13 has two of its three external electrons removed by a sufficiently intense spark discharge, a structure of 11 electrons like that of sodium is again obtained and its spectrum likewise is similar to that of sodium. Owing to two extra positive charges on its nucleus, however, its spectral series are of even higher frequency than those of singly ionized magnesium. The spectra of atoms which have lost one or more of their external electrons are often referred to as spark or enhanced spectra. A considerable amount of work has been done in the study of the spark spectra of various elements, particularly by Millikan and Bowen at the California Institute of Technology. In the very violent conditions which exist in stars, atoms of various degrees of ionization are common, and in all cases their spectra are found to check with those observed in the laboratory.

1033. Further Remarks on Atomic Spectra. Complicated as spectra may seem from this discussion, there are still further complications not mentioned. The various lines referred to are practically all of them found to be doublets, triplets or to be made up of even more than three lines very close together which can be separated with a spectroscope of high resolving power. The discussion of these complexities is given in more advanced treatises such as those by Darrow and Sommerfeld cited at the end of this chapter.

Atoms with many outside electrons such as those in columns IV to VII of the periodic table (§ 1028) have spectra whose complexity is terrific when resolved by a powerful spectroscope. It is remarkable that so much progress has already been made in unraveling them. A considerable portion of modern research is devoted to the study and analysis of such spectra in order to throw more light on atomic structure, and the nature of the emission and absorption of radiation.

DIRECT MEASUREMENT OF ENERGY LEVELS

1034. Direct Measurement of Energy Levels; Excited Atoms; Radiation Potentials. Important work has been done in recent years on the direct measurement of energy states of various atoms, including their ionization potentials. The pioneer work in this field was done by Franck and Hertz in Germany, and was published by them in 1914. As an illustration of the type of method used of which there are many variations, one used by K. T. Compton of Princeton University will be described which is a modification of a method first used by the German physicist, Lenard. This is shown in the diagram of figure 633.
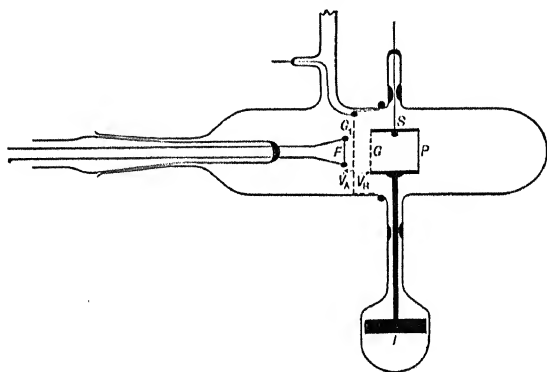
FIG. 633. Compton's apparatus

It consists of a tungsten filament F from which electrons are emitted, a wire gauze G_1, a receiving electrode which consists of a box with the wire gauze G in front of it. This box can be turned about the swivel S by means of the iron armature I. The three electrodes are contained in a glass vessel, into which is placed the gas to be studied at the required pressure. When the voltage of the gauze G_1 is raised an amount V_A, above that of the filament F by a battery, or some other suitable means, the electrons from the filament F are attracted toward it and most of them pass through it with a velocity determined by the voltage V_A. This velocity is readily calculated from the relation:

$$\tfrac{1}{2}mv^2 = eV_A. \tag{11}$$

This equation is seen to be the same as (3) of § 1018. The electron in such a case as this is said for abbreviation to have a velocity of V_A equivalent volts. It is of course necessary that the pressure of the gas

used be sufficiently low so that the motion of the electrons is not interfered with by impacts with molecules before G is reached. The distance which an individual electron may travel in such a gas increases as the pressure is decreased, the pressure required in this case being less than 1 mm. of mercury.

The receiving electrode is maintained at a voltage slightly lower than the filament such that V_R is greater than V_A. This means that the electrons passing through G_1 with a velocity of V_A equivalent volts, lose all of their velocity before reaching G and unless some other action is at work the electrometer, which is connected to the receiving electrode, shows no deflection because it receives no electrons.

Suppose, however, that the electrons passing through G_1 have sufficient velocity so that atoms of the gas such as helium are *excited* by them, that is, one of the external electrons of some of the gas atoms is knocked out of its normal position by an impact from one of the oncoming electrons, and lifted to a higher energy level, without being actually removed from the atom. After remaining in the excited state for an instant, the excited atoms assume their normal state again owing to their displaced electrons dropping back into their normal positions, in each case *emitting a quantum of radiant energy*. The size of this quantum (and also the radiation frequency) is, as before, determined by the difference between the energy of the atom in its excited state and the atom in its normal state. But as soon as radiation is produced, photoelectrons (§ 1016) are emitted from the receiving electrode, the rate of this emission being recorded by the quadrant electrometer. Evidently, as soon as the electrometer begins to register photoelectrons while V_A is being increased, the voltage V_A has just reached a value sufficient to produce excited atoms, and a direct measure of the corresponding energy state is obtained in equivalent volts above the neutral state. Potentials corresponding to different energy states below the ionizing potential registered in this way are called *radiating potentials* since they are recorded by the radiation produced.

Although measurement shows that an electron may remain in such an excited state only about 10^{-8} seconds before dropping back into the neutral state again, this time is probably very long indeed compared with that required for the actual lifting out process or the dropping back process.

1035. Measurement of Ionization Potentials; Discussion of Curves Obtained. It is also possible, by this method of Compton's, to distinguish between radiation potentials and ionization potentials. The latter are measured by values of V_A such that the electrons from the filament acquire just sufficient energy to knock external electrons

entirely out of atoms they collide with, thus ionizing them, instead of
lifting them only partially out of the atoms to intermediate energy
levels.   When ionization is produced, positive ions appear.   These
positive ions are attracted instead of repelled by the receiving elec-
trode, so that the quadrant electrometer at once registers a charging
current.   Since the emission of electrons by the photoelectric action of
radiation from the gas at a radiating potential also leaves the elec-
trometer positively charged some device is necessary to distinguish
between the radiation and ionizing potentials.

   This is accomplished in Compton's apparatus by first measuring the
charging current with the gauze G of the receiving electrode toward
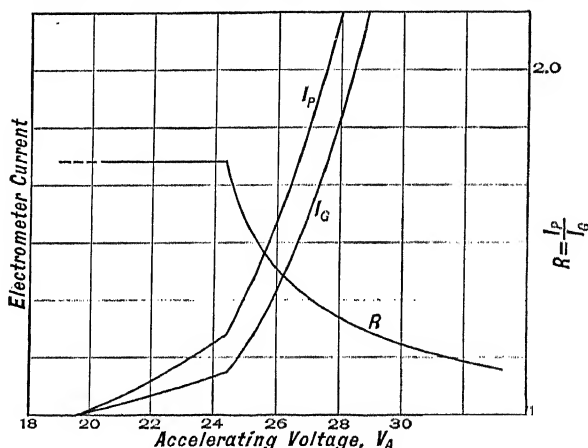
FIG. 634.   Curves for helium obtained by Compton's method

the filament and then revolving this electrode about on its swivel so
that the side P of the box is toward the filament.

   Positive ions are evidently attracted to the receiving electrode as
fast as they appear whichever way this electrode faces.   But the case is
different with photoelectrons.   When radiation is produced by excited
atoms the number of photoelectrons emitted by the receiving electrode
is less when the radiation passes through the gauze G into the box than
when it strikes the side P directly, as it does when this side is toward
the filament.   This is because most of the photoelectrons emitted in-
side of the box are trapped there and received by it again, only those
emitted from the gauze permanently escaping, whereas they all escape
when the radiation strikes P directly.

   Figure 634 shows typical forms of electrometer current curves as de-
pending upon the *accelerating* voltage V_A which determines the veloc-

ity of the electrons. No current appears until a voltage of about 19.75 is reached when current begins to be registered by the electrometer. The curve marked I_P shows the current obtained with the closed side of the box toward the filament and that marked I_G the current observed with the gauze-covered side of the box toward the filament. Since I_P is greater than I_G at least part of the effect is due to a radiation potential having been reached. To separate a possible current due to ionization which may be superimposed upon that due to the radiation potential, the curve R is plotted, which gives the ratio of the current I_P to the current I_G. It is noticed that a sudden drop in this curve begins at about 25 volts and it is here that the atoms begin to be ionized. As long as the ratio of the I_P to I_G is a constant, the effect is in all probability due to photoelectric action, or to radiation from excited atoms only, but as soon as this ratio begins to drop this change is explained as due to an increment of current superimposed upon P and G in *equal* amounts. This is exactly the way a positive ion current would be recorded since it registers the same for either position of the receiving electrode as described above.

We have thus recorded a radiation potential at 19.75 volts and an ionizing potential at about 25 volts. The gas used in this case was helium. Helium is a monatomic gas and therefore one of the easiest to study in this way. The measurement of radiation potentials and ionization potentials of molecules containing more than one atom per molecule presents greater difficulties.

It is important that the electrodes used be of the same metal, as far as possible, so that the opposing contact potentials of the electrodes (§ 778) shall cancel out, and thus do not have to be taken account of.

By methods such as this the potentials of various energy states including the ionizing potential of many types of atoms have been directly measured giving important evidence of the existence of stationary energy states and of the absorption and emission of radiation in quanta of size determined by these energy states. Below is given a table of some radiation and ionization potentials of a few different types of atoms obtained by such measurements.

RADIATION AND IONIZATION POTENTIALS (after *K. K. Darrow*)

	HE	NE	NA	CS	MG	HG
Energy value of the normal state.	0	0	0	0	0	0
First excited state	19.75	16.65	2.1	1.45	2.7	4.66
Other excited states	20.55	18.45			4.4	4.86
						5.43
						6.7
Ionized atom..................	24.5	21.5	5.12	3.9	7.6	10.4

1036. Elastic and Inelastic Collisions; Collisions of the Second Kind. In work of this kind the terms *elastic* and *inelastic* collisions are frequently used and deserve some explanation. When gas molecules or atoms are bombarded by a stream of electrons of comparatively low velocity, say of 10 equivalent volts, for the case of helium, the electrons bound back and forth against them like so many perfectly elastic spheres, all of the energy being retained as kinetic energy of the electrons and also of the atoms of the gas, with which some of the kinetic energy is exchanged. The collisions between the molecules of gas are perfectly *elastic*. As soon as the velocity of the electron stream is sufficiently high, however, to produce excited atoms, each electron which produces by its impact an excited atom delivers part or all of its velocity energy to the atom in lifting one of its electrons to a higher energy level. Instead of rebounding, it strikes it a more or less dead blow as though the atom were a sphere of putty. This type of collision is evidently *inelastic*. The appearance of excited and ionized atoms in a gas so bombarded is thus the result of inelastic electron collisions.

A very interesting type of collision sometimes observed in this work is called a *collision of the second kind*. In this case an atom already excited goes back to its normal state, at the same time shooting out a neighboring electron which happens to be within range at the right time, delivering its quantum of energy to the electron instead of emitting the quantum as radiant energy.

THE QUANTUM THEORY OF MOLECULAR VIBRATIONS AND SPECTRA

1037. Band Spectra. Band spectra is a term used where the lines of various spectral series under low resolving power are broadened out into more or less continuous bands. They are sometimes known as fluted spectra (§ 938) or may be so abundant as to appear like a continuous spectrum. Whereas line spectra are mainly due to processes within individual atoms, band spectra are produced by molecules or groups of molecules excited by electric discharge or by high temperature, and continuous spectra by solids and liquids at high temperature.

. It has been seen that even atoms by themselves may have very complex spectra, but these complexities greatly multiply when atoms are bound together to form molecules. Instead of individual lines, multitudes of lines appear which are seen as bands which often cannot be resolved at all. Just as for atoms, however, every individual line or element of a band is a measure of the emission of a definite quantum of energy, and, as before, the size of this quantum is found by multiplying the frequency ν of the line by Planck's quantum of action h, giving $h\nu$.

The extremely large number of lines of molecular spectra thus indicates the possibility of many times more energy states than for atoms, since, as in the case of atoms, molecular spectrum lines arise from the emission of quanta corresponding to energy changes from higher to lower energy levels.

The complete spectrum appears because large numbers of molecules at the same time are each undergoing a particular change from a higher to a lower energy level, and furthermore the number of such changes possible per molecule in one second is enormous, so that by the law of probability every individual line makes its appearance, the more probable ones being the brighter.

1038. Quantizing of Molecular Vibrations. The process of assigning definite energy levels to atoms as determined by definite electron orbits is spoken of as the *quantizing* of the electron motion. The transfer of the electron from one energy level to another is attended by the absorption or emission of a quantum of energy, depending upon whether the transfer is from a lower state of energy to a higher or vice versa. In like manner Bohr and his followers have quantized the motion of whole atoms within molecules and indeed the motions of the molecules themselves. To take a simple case, a symmetrical diatomic molecule may be thought of as shaped like a dumbbell. The two atoms which are held together by their mutual attraction may vibrate to and from each other along the axis of the dumbbell. The dumbbell as a whole may revolve about an axis perpendicular to its own axis, but not about one coinciding with it, since, owing to its symmetry and circular cross section, impacts cannot excite a rotation about this axis. It is clear that the more rapid this vibratory motion or this rotational motion the greater the energy of the atom. These two forms of motion are each quantized so that again only certain particular vibration frequencies and rotational speeds are possible, corresponding to series of definite energy states, as before. These are determined by a logical development of the quantum theory, best given by Sommerfeld, cited at the end of the chapter. When these new energy states are combined with those of the atoms themselves, which make up the molecule, the sum total of possible energy changes becomes enormous. It is indeed remarkable that in many cases studied the lines predicted by this theory are actually observed. Molecular hydrogen and the gas HCl are good examples.

1039. The Quantum Theory and Energy Distribution in the Spectrum of a Black Body. It has already been briefly noted in § 484 in the chapter on radiation of heat that the first suggestion of the absorption and emission of radiation in quanta was given by Planck in order to account for the way black body radiant energy is distributed

among wave lengths, as shown by the curves of figure 253. Ideal black body radiation (§ 483) has a continuous spectrum as shown from curves obtained by experiment, such as those of figure 253.

Now if it is attempted to derive a curve of energy distribution among wave lengths for black body radiation on the basis of the electromagnetic theory it has been found that the shorter the wave lengths the greater the energy, so that the major part of the energy is concentrated in a narrow band including only very high frequencies. In fact if any frequency whatsoever is permitted to exist, however high, the total energy of the spectrum becomes infinite, which is an impossible result. On either basis the energy distribution curves so obtained do not coincide with those of figure 253, but show a maximum at the shortest wave length present, at the left-hand side of the figure, instead of showing a definite maximum and then falling off to zero gradually.

On the basis of the quantum theory, however, the size of the quanta being $h\nu$, they become very large at the high frequencies and it requires so much energy to produce sufficient excitation of the atoms that the emission of these quanta becomes less and less probable the higher the frequency so that the greater energy they individually contribute is soon offset by the smaller number of them produced. This accounts for the energy curves reaching a maximum and then falling off again for the very high frequencies or short wave lengths. If the temperature of the body is raised, however, the atoms are more violently excited resulting in the emission of more high frequency quanta. It thus follows that the higher the temperature the farther toward the short waves the point of maximum energy is displaced as shown in figure 253.

In his analysis in which he proposed the idea of absorption and emission in quanta, Planck expressed his ideas in mathematical form and the energy distribution curves obtained by him show good agreement with those obtained by experiment for a black body.

For a good discussion of Planck's development of the quantum theory, see that of Lewis cited at the end of the chapter.

1040. The Quantum Theory of Specific Heats of Solids. It had long been known that the specific heats (§ 409) of certain substances vary somewhat with temperature, particularly at very low temperatures where a great decrease is observed, the specific heat becoming vanishingly small toward the absolute zero.

According to the law of Dulong and Petit (§ 416) the product of the specific heats of the various elements into their atomic weights, that is, their atomic heats, are approximately a constant, which is equal to about 6 gram-calories. This value for the atomic heat of an element was satisfactorily explained on the basis of the kinetic theory,

making certain assumptions as to the molecular motions, but the kinetic theory fails to explain the decrease of atomic heat with decreasing temperature (§ 415).

Einstein, in the year 1907, was the first to point out that Planck's quantum theory would afford an explanation of this decrease of atomic heat with temperature, and he derived a mathematical expression based upon Planck's theory of quanta which gave the atomic heat of a substance as depending upon its temperature, and which showed a decrease with decrease of temperature which checked experimental measurements very well. Figure 635 shows the calculated curve of
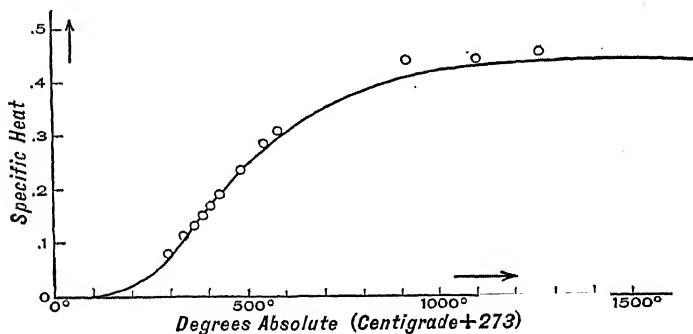
FIG. 635. Variation of specific heat of diamond, with temperature

the temperature variation of the specific heat of diamond compared with points obtained by measurement, according to Einstein. The agreement in this case is excellent. It is not so good in other cases, however.

Other attempts have been made subsequently to derive a similar law of specific heat variation on the basis of the quantum theory. The most successful is that of the Swiss physicist, Debye, published in 1912, whose method is superior to that of Einstein in that it takes account of a whole series of possible harmonic vibration frequencies as making up the thermal agitation within the substance instead of one characteristic molecular vibration frequency throughout as assumed by Einstein. The law derived by Debye requires a somewhat more complicated mathematical expression than that of Einstein, but it shows a decidedly better agreement with measurements made on many elements.

For a fuller discussion of this subject see Lewis, cited at the end of the chapter.

This application of Planck's quantum theory to the specific heats of solids is particularly important because through this work the possibilities of the theory became known and appreciated to a much greater degree than before, and the way was paved for the remarkable work of Bohr and his followers.

THE QUANTUM THEORY OF X–RAYS

1041. Moseley's Experiments. In § 1020 a brief discussion of X-rays was given since the emission of X-rays affords a good illustration of an inverse photoelectric effect. We are now in a position to give a more complete discussion of X-rays in the light of Bohr's development of the quantum theory. The famous experiments of the English physicist, Moseley, performed at the Cavendish Laboratory, and published in 1913, will be described next as these experiments were the first to clearly bring out the basic laws of X-ray spectral series.

Moseley obtained powerful beams of X-rays from a succession of different elements by using each, in turn, as the anode of his X-ray tube. Electrons, driven against these anodes by voltages of high intensities excited the X-rays characteristic of the anode materials. The elements examined by him covered the range of atomic numbers from 20 to 30, that is, from calcium to zinc, except for scandium, inclusive (see § 1028 and Fig. 636).

FIG. 636. Moseley's photographs

The beams of X-rays from these materials were each analyzed by an X-ray spectrometer similar to that shown in figure 601, and photographs are shown in figure 636, as given in Moseley's original paper in 1913. Moseley arranged these photographs in the figure according to wave length, so that spectrum lines of shorter wave length are at the left of the figure, and those of longer wave length at the right. By so doing, he was able to bring out and to formulate a very remarkable relation between the wave frequencies and the atomic numbers of the

elements which produced them, the wave frequency being simply $\frac{c}{\lambda}$ where c is the velocity of light and λ the wave length.

He found that the *frequency corresponding to a characteristic spectrum line of any element is proportional to the square of its atomic number.* For instance, in the photographs of figure 636, each element shows two lines, one of higher and one of lower frequency. Either the higher frequency series of lines from one element to the other, or the lower frequency ones are found to obey Moseley's law of frequencies proportional to the atomic numbers squared.

Now if the frequencies are proportional to the squares of the atomic numbers, their square roots should be proportional to the atomic numbers, in other words a curve of the square roots of the frequencies plotted against atomic numbers should be a straight line. This has been found to be the case over the entire range of known elements to a close approximation, except for elements of low atomic number. The truth of the relation is shown in figure 637 from elements of atomic numbers 11 to 60.

This discovery marked an epoch in the subject of X-ray spectroscopy. It was the first and a fundamental step in the systematizing of X-ray spectra.

This work of Moseley's is especially remarkable because he was a young man of only twenty-six at the time, and it is indeed a sad thing that he soon after lost his life in the Great War.

1042. Characteristic X-ray Spectra. The pairs of spectrum lines observed by Moseley, and shown in figure 636 represent only the highest X-ray frequency lines of the elements, and are called the K lines. It is now known that other lines exist, of lower frequency, which are called the L and M lines, and even others are possible. Furthermore, the K lines shown are now known to be only part of a whole series of K lines which exist for each element, although the others are fainter and nearer together and thus harder to observe. The same is true of the L series and the M series as they are called. In fact a close analysis shows that even the two highest frequency K lines are really doublets, and lines in the other series are even more complex. A more complete analysis and classification of these lines requires an intimate knowledge of energy changes within the atom, which will be introduced in an elementary way in some of the following sections.

Figure 638 shows some of the L and M frequencies as well as the K plotted against atomic numbers as in figure 637.

Figure 639 shows the same thing as figure 638, except the lines are shown depending upon wave length instead of the square root of the frequencies corresponding to these wave lengths.

The K lines observed by Moseley were the first to be studied because they are of such wave length as to be observed easily by a spectrometer employing crystals as diffraction gratings. Wave lengths greater than twice the atomic spacing in the crystal used cannot be measured at all by the crystal method while those much shorter than the atomic
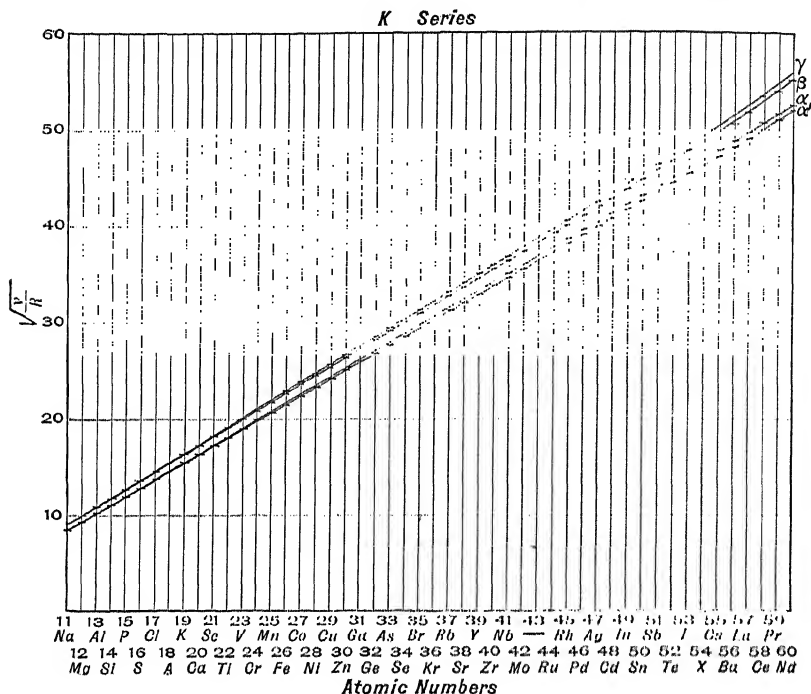
Fig. 637. Curve illustrating Moseley's law (after *Sommerfeld*)

spacing are diffracted at such angles that θ becomes almost $90°$ and very little dispersion is possible.

1043. Further Remarks on X-ray Spectra. It was believed for many years that X-ray spectra and visible spectra (the term visible spectra as here used means spectra whose frequencies are of the same order of magnitude as those of visible spectra) were two entirely distinct phenomena, and governed by different laws, but in the light of recent discoveries the barrier has gradually broken down, and it is now known that they represent the same type of radiation except that X-rays are of a far higher frequency than visible spectra. Further-

more X-rays as well as visible spectra are governed by Bohr's fundamental frequency condition $h\nu = W_1 - W_2$ where W_1 and W_2 represent the energy in two of a whole series of possible states of the atom.

Since X-rays have very high frequencies compared with visible light the corresponding quanta are relatively large packets of energy. This
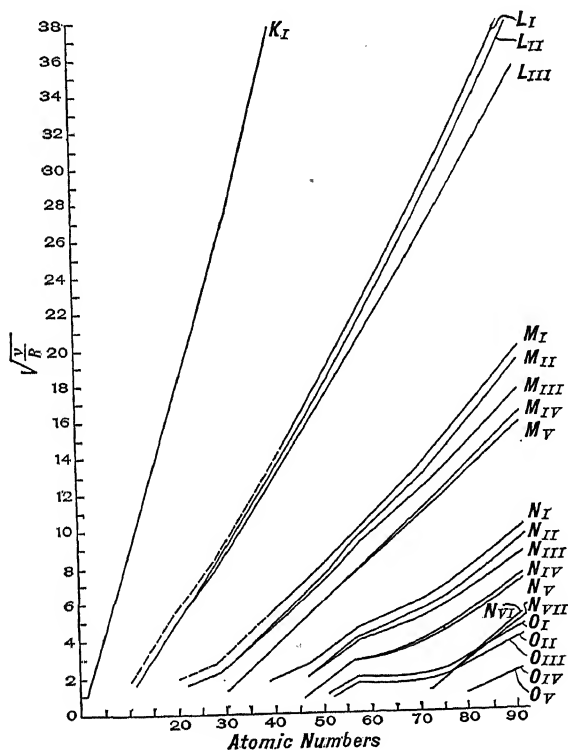
FIG. 638. (*Sommerfeld*)

in turn requires that the energy shift responsible for each quantum emitted be also very large. This is only possible on the supposition that very deep-lying electrons are knocked or lifted from their normal positions by the cathode ray bombardment, the radiation being emitted when an electron drops back again to these very low energy levels. The energy level of the displaced electron must necessarily be very low compared with the ionization energy state in order that the

energy shift shall be large enough to produce the large sized quanta of X-rays. For this reason, elements of low atomic number which are built up of comparatively few electrons, none of which lie deep within the electron structure, have not the power to produce high frequency X-rays. A metal like tungsten, however, of atomic number 74 may emit X-rays of high frequency, as its deep-lying electrons are far below the outer shell of the electron structure. Furthermore the highest frequency X-rays obtainable from an element, that is, the K radiation
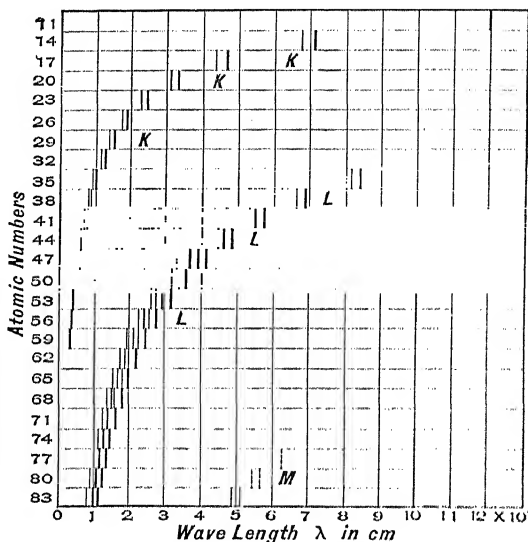
FIG. 639. K and L lines for series of elements plotted against wave length

(§ 1042), is now known to be a consequence of lifting the very deepest-lying electron out of the atom, the L radiation the next deepest, etc., the series continuing up to the visible spectrum produced by the outside valence electrons. On this basis the Lyman series of hydrogen (§ 1023) is the hydrogen K radiation, the Balmer series the hydrogen L radiation, etc.

Since the hydrogen atom has only one electron in its external structure its K series is that produced by the falling of this electron into its lowest or normal level when displaced by excitation, since this one electron is evidently its lowest possible electron. This gives the Lyman series. Similar considerations show that its L series must correspond

to the Balmer series. For elements of larger atomic number the lowest lying electron is evidently deeper within its structure, and the K radiation of correspondingly higher frequency.

An important difference between the production of X-ray spectra and visible spectra is that no X-ray spectrum lines appear at all until the excitation is sufficiently great to entirely remove a K or an L or an M electron from the atom, and then the whole K or L or M series appears at once, depending on which of the 3 electrons is removed. It is not possible to obtain the radiation of first one then the next higher, then the next higher line of these series, etc., one after another, by excitation of the right intensity, as in the case of visible spectra. This is doubtless because visible spectra are produced by the displacement of the outermost electrons of the atom to unoccupied orbits (§ 1027) which contain no electrons, in the normal state of the atom, whereas the important lines of X-ray spectra arise from electrons dropping from energy levels normally occupied by electrons. Therefore before a readjustment can take place involving the whole series of energy levels down to the deep-lying electron this electron must be entirely removed from the atom.

It is true that it may be sometimes only removed as far as some outside unoccupied orbit, but the differences in energy levels of the outside unoccupied orbits are so small compared with the internal X-ray energy level differences, that any observation of this effect is entirely lost because the part of the X-ray spectral series due to the unoccupied orbits must be all concentrated right at the limit of the series, that is, almost at the point corresponding to complete ionization.

1044. X-ray Continuous Spectra. A very noticeable characteristic of X-ray spectra (and to some extent in visible spectra) is the presence of a continuous background of less intensity than the individual spectrum lines which may continue up to frequencies considerably higher than the limit of even the K series. This continuous spectrum shows a definite high frequency limit, however, beyond which it disappears. This limit is found to depend upon the voltage applied to the tube in driving the electrons against the target which emits the X-rays, and to accurately obey Einstein's law as discussed in § 1018. This is thus a good example of an inverse photoelectric action, by which the velocity energy of some of the electrons striking the target is transformed into quanta which are radiated out along with the characteristic radiation. The largest quanta obtainable are evidently fixed by the total voltage applied to the tube. The continuous series extending below this limit can be explained as also due to an inverse photoelectric action arising from electrons whose velocities have been decreased in various amounts due to their having delivered part of

their initial energy to the cathode in various ways such as inter-
ference with other electrons, etc. The curves of Duane and Hunt
(Fig. 629) show that the intensity of radiation of a given wave length
rapidly increases with distance from the edge of this continuous spec-
trum toward longer wave lengths. The curves of Duane and Hunt,
however, were obtained by selecting a fixed wave length of definite
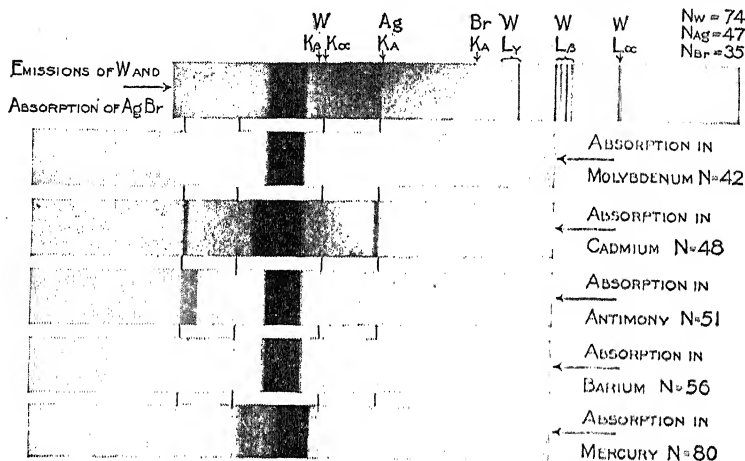
FIG. 640.   X-ray absorption spectra, K series.   (Photograph by
De Broglie from *The Electron* by Millikan)

frequency by means of an X-ray spectrometer and lowering the tube
voltage until the high frequency limit reached that frequency instead
of maintaining a constant voltage on the tube and observing the in-
tensity of the radiation for increasing wave lengths until the high
frequency limit was reached. The law of falling off of intensity,
however, should be nearly the same in either case.

1045. X-ray Absorption Spectra. The absorption spectra of X-rays
also show a marked difference from those of visible spectra. In the
case of absorption spectra of the visible type it will be recalled that the
various spectrum lines become reversed so that they appear as dark
lines such as the Fraunhofer lines in the spectrum of the sun (§ 932).
When X-rays are passed through a layer of absorbing material, how-
ever, broad absorption bands appear as well as individual dark lines
characteristic of the absorbing material. This is very well illustrated
in the photograph of figure 640 which was taken by De Broglie in
Paris in 1910.

The upper band shows the spectrum of tungsten obtained from the
tungsten anode of an X-ray tube bombarded by electrons. Wave

lengths decrease and frequencies increase toward the left until the very dark band is reached which is made by the undiffracted X-ray beam opposite the spectrometer slit. To the left of this the spectra repeat again in the opposite direction, that is, with increasing wave length and decreasing frequency. The tungsten spectrum is seen to show various K and L lines, also the continuous spectrum is seen to extend somewhat to the left of the highest frequency K radiation, indicating the voltage applied to the tube from Einstein's photoelectric equation (§ 1018). The second band shows the same spectrum, but photographed in this case after being transmitted through a thin sheet of molybdenum. It is seen that beyond a definite frequency limit strong absorption begins, this being indicated by the light band since absorption screens the transmission of light. This limit is the high frequency limit of the K radiation. Since the frequency of the radiation below this limit is insufficient to lift the K electrons up to the surface of the atom because the corresponding quanta are not sufficiently large, this radiation cannot be absorbed in this way, so that much more of it passes through the molybdenum sheet than for radiation of frequency just above the K limit. This latter radiation has the power to lift the K electrons entirely out of the atom and thus excite the whole K series and thereby loses much of its energy, which results in strong absorption. These *secondary X-rays* excited in the molybdenum appear largely at the surface towards the X-ray source away from the photographic plate and are thus radiated back without affecting it. The absorption shows as a broad band because frequencies considerably higher than the K limit also excite secondary K radiation and are also absorbed, but to a less extent than frequencies immediately above the K limit because the former have more penetrating power than the latter due to their higher frequency. For this reason the K absorption band begins suddenly with strong absorption at the K limit and gradually fades out with higher frequencies. Absorption bands also appear at the high frequency limit of the L series which fade out with higher frequencies and disappear before the region of the K radiation is reached. The same is true of the M high frequency limit, etc. This fading out of the absorption band with higher frequency shows well for the K absorption band of antimony in the fourth photograph from the top of the figure.

This figure has other interesting points. In the tungsten spectrum shown in the upper band where no absorbing medium is interposed the K absorption bands of both silver and bromine appear, but as dark instead of light bands. These are beyond doubt due to the silver bromide in the coating of the photographic film. They appear dark because the photographic layer absorbs these frequencies more strongly

causing it to be more strongly affected than otherwise, thus magnifying the intensity of the incident radiation within these frequency limits. This effect shows through the absorbing material to some extent as observed in the five lower photographs.

This important X-ray phenomenon was first discovered by the English physicist, Barkla, some years before its full significance was understood.

It can be said of X-ray absorption in general, that the greater the atomic number of the absorbing layer the more powerfully does it absorb, the absorption increasing very rapidly indeed with atomic number. Lead screens prove very effective in protecting X-ray operators from stray X-rays which may produce dangerous burns without them. Glass containing a high percentage of lead is used in some applications, for this reason.

1046. X-rays from Radioactive Transformations and Cosmic X-rays. It will be recalled from the introductory discussion of radioactive radiations (§ 787) that the gamma rays are nothing other than very "hard" or high frequency X-rays emitted during the breakdown of the atoms of radioactive substances. These X-rays show higher frequencies than have thus far been produced by laboratory methods, being radiated in correspondingly large sized quanta. In recent years very definite evidence has been obtained of X-rays of extraordinary hardness which come to the earth in extremely small amounts from interstellar space. These *cosmic rays* as they have been called have been detected by various investigators through the observation of a slight ionizing effect detected through the discharge of an inclosed electroscope (§ 537) which no ordinary precautions in the way of screening could eliminate. This suggested the possibility of the presence in space of X-ray radiations, of penetrating power far greater even than that of gamma rays, which cause the discharge of the screened electroscope by a slight ionizing effect which persists in spite of the screening.

The first serious study of this possibility was made by the Swiss physicist, Gockel, who made observations from a balloon which showed that at a height of 4500 meters above the earth the effect was still present at its full intensity. Other investigators subsequently obtained evidence of an increase of the intensity of this radiation with altitude.

The most conclusive and complete study of this phenomenon has been made by Millikan and his collaborators at the California Institute of Technology. Observations made below the surface of snow-fed lakes, at high altitudes, showed that the effect could be detected at a depth of over 50 feet and that the effect fell off with increasing depth

at a rate to be expected from very hard X-rays. Furthermore these rays *increase* in intensity with altitude and are thus doubtless of cosmic origin. It is not known how they arise, except that the atomic energy changes responsible for them are much greater than anything encountered on the earth. This radiation is emitted in quanta over thirty times larger than the largest quanta known from any other X-ray radiation. The probable wave lengths of these cosmic rays is only about .0005 Å (one Ångstrom = 10^{-8} cm.).

1047. The Compton Effect. No discussion of X-rays and the quantum theory can be complete without some mention of a remarkable effect discovered by A. H. Compton of Chicago University and studied particularly by him and by Duane of Harvard University. It was afterwards explained by Compton on the basis of the quantum theory, thus being named after him the *Compton effect.*

When X-rays of a particular wave length strike a piece of matter such as a plate of graphite, secondary X-rays are found to be scattered in all directions. If rays scattered in some particular direction, say 90 degrees from the direction of the incident beam, are selected and analyzed by an X-ray spectroscope the selected beam will be found to be made up of a series of wave lengths with two prominent maxima as shown in figure 641 showing that two definite wave lengths are present in far greater abundance than those of slightly different frequencies. These curves were obtained from measurements made by Compton. The original

Fig. 641. Curves showing Compton effect

beam, which was used, was obtained from an X-ray tube with a molybdenum target. This was analyzed by a spectrometer and the K line alone of the K series was used to excite the secondary radiation from the plate of graphite, this secondary scattered radiation being examined as described above by a second spectrometer for rays scattered at the three angles 45 degrees, 90 degrees and 135 degrees. The remarkable fact which appears is that in each case, besides the original frequency of the K line appearing in the scattered

radiation shown in each case by the left-hand maximum, a somewhat longer wave length is present which diverges more from the original frequency the greater the angle of scattering. The crystal used in the spectrometer was calcite, and from the angle at which the diffraction takes place from the calcite and the atomic spacing of the calcite, the wave lengths are easily found according to equation (1) of § 978. This angle is seen to be measured along the horizontal coördinate in the figure.

Compton has given the following very ingenious explanation according to which at each angle of scattering a particular frequency of scattered radiation should appear which is lower than that of the incident beam. He assumes that certain quanta in the radiation incident on the scattering substance such as $h\nu$ in figure 642 come in contact with a free or practically free electron, delivering an impact such that the electron goes in one direction shown by the lower
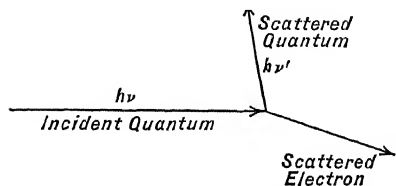
FIG. 642. Impact of quantum with electron

arrow and a quantum goes out in another direction shown by the upper arrow. This is the scattered quantum, of somewhat smaller size and lower frequency than the original. In order to analyze such an impact, however, the quantum must be considered to have momentum as well as kinetic energy, the latter being taken as simply $h\nu$. Now it is well known that light incident upon an absorbing surface exerts a definite pressure $\dfrac{I}{c}$ where I is the energy of the beam absorbed by the surface per second, and c is the velocity of light. It can be shown on this basis that, according to the electromagnetic theory, the momentum delivered by a quantum equals $\dfrac{h\nu}{c}$. This is the value of the momentum of a quantum of radiant energy taken by Compton. For the electron, moving with a velocity v, according to elementary mechanics its kinetic energy is $\frac{1}{2}mv^2$ and its momentum mv where m is its mass.

Thus when the quantum strikes the electron it delivers to it a certain amount of momentum and energy, itself losing the kinetic energy delivered and also the momentum component delivered.

Using the above fundamentals as a basis, Compton worked out a simple mathematical expression which, for any angle of scattering, gives the size of the scattered quantum $h\nu'$ and thus the frequency ν'. It is found that similar measurements made by many other observers

as well as those of Compton himself give values for ν' which check the theoretical value within experimental error. The Compton effect is of very great interest to students of the quantum theory, because it shows directly the action of a quantum upon an individual electron.

Compton has obtained a still further check on this theory from the observation of the paths of electrons scattered from a beam of X-rays passed through a gas. These paths are photographed by using the method of C. T. R. Wilson described in § 789. A large number of photographs were taken, and a few of them caught an electron in the act of being deflected.

1048. The Paradox of Modern Physics. How can radiation be a wave phenomenon and at the same time be emitted in darts or packets called quanta? Evidence that radiation travels in waves is abundant, as can be seen by referring back to the wave theory of light. Even very short X-rays are clearly transmitted as waves. And yet they seem unable to be received by matter or emitted by it except in bundles or packets of minute size called quanta. How is it possible for radiant energy in the form of waves distributed through a considerable volume of space to suddenly concentrate itself upon a single electron as a quantum? And yet the quantum itself is defined in terms of frequency measured on the basis of the wave theory. The situation presents a curious paradox in which radiation phenomena are explained on the basis of two radically different theories which are curiously intertwined with each other and for each of which indisputable experimental evidence is at hand.

1049. Foreshadowings of the Future. Meanwhile a large group of able theoretical physicists, among whom Germany is well represented, and the Danish physicist, Bohr, is an outstanding figure, have been working at the problem and developing an entirely new viewpoint which shows promise of a far better understanding of the difficulty. Never before in the history of physics has such a concentrated organized effort been made and it is noteworthy that those who are taking part are mostly comparatively young men. The problem is of such a nature that unusual powers of mind are required to handle it, with which comparatively few are endowed. ·

At the same time, experimental work continues, as the test of the truth of any physical hypothesis or theory must lie ultimately in experimental verification. Some outstanding results have recently been obtained by Davison and Germer of the Bell Telephone Laboratories in the study of the way a beam of electrons is "reflected" from a crystal of nickel. The stream of electrons was directed normally on the crystal, the velocity of impact being determined by the applied voltage through which the electrons are made to drop according to

equation (3). The entire apparatus was sealed in a vacuum tight chamber. The very remarkable result was obtained that the electrons were "reflected" much more strongly at some angles than at others, a complete survey showing a reflection pattern similar to that of the Laue pattern (Fig. 598) obtained from X-ray "reflection." This indicates that an electron stream in some way has the power to act as though it were a wave phenomenon. It is of great interest that recent theory has suggested the possibility of such a result.

1050. Astronomical Applications of Modern Physics. The most remarkable advance has taken place in our understanding of stellar phenomena during the past decade, largely as the result of the analysis of the radiation coming to us from stellar bodies on the basis of Bohr's theory and its developments. In the case of the hotter stars, the atomic activity is so intense that it is now recognized that the atoms of some of the elements are stripped of several of their outermost electrons much of the time, thus emitting spectrum lines characteristic of these stripped atoms. An extreme case of this sort is that of the companion of Sirius which has been found to have the extraordinary density of about 60,000 times that of water, that is, one cubic inch of material from this star would weigh about a ton on the earth. It is believed that the atoms which make up this star have lost most of their external comparatively light electron structure and have become crowded together so that the relatively massive atomic nuclei (§ 807) are 30 or 40 times closer together than in an ordinary heavy material.

According to calculations made by the English mathematical astronomer, Eddington, the internal temperature of most stars is of the order of 40,000,000 degrees C. This inconceivably high temperature is attended by radiation so intense that the outward pressure from it (§ 1008) may actually balance the enormous gravitational pressure in the case of the largest stars, and so prevent their existing above a certain mass. This conclusion seems to be verified by observation, according to which stars are never found above a mass, say, of about 30 times that of the sun, nor indeed are they often found of mass much smaller than the sun.

This whole subject is of very great interest and is beautifully presented in a little book by Eddington called *Stars and Atoms*.

REFERENCES

K. K. DARROW: *Introduction to Contemporary Physics.*

W. C. McC. LEWIS: *A System of Physical Chemistry*, Vol. III: *Quantum Theory.*

A. SOMMERFELD: *Atomic Structure and Spectral Lines.*

A. S. EDDINGTON: *Stars and Atoms.*

APPENDIX I

CARNOT'S THEOREM AND THE ABSOLUTE SCALE OF TEMPERATURE

1. Work of Compression or Expansion. When steam is admitted to a cylinder C and the piston P is pushed back through a distance x, the force exerted by the steam against the piston is pA, where p represents the pressure of the steam and A is the area of the piston head, and the work done is pAx, or

$$W = pAx.$$

But as the piston moves through a distance x the volume of the cylinder is enlarged by the amount xA, so that, representing the increase in volume by v, we have

$$W = pv.$$
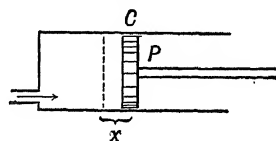
FIG. 643

That is, *when any substance while exerting a pressure p expands in volume by an amount v, it does an amount of work equal to the product pv.*

2. Work on the Pressure-volume Diagram. The work of compression or expansion is conveniently represented by an area on the *pressure-volume* diagram. For let volumes be represented by abscissas, or distances measured to the right of the line OP according to the scale at the bottom, and let pressures be measured by *ordinates*, or distances measured upward from the base line OV according to the scale at the side; then if the pressure remains constant at, say, 40 lbs. to the inch, while the piston moves enlarging the space from zero to 5 cu. ft., the series of states will be represented by the successive positions of a point R as it moves from A to B (Fig. 644). The work against the piston which we have seen to be equal to the product of the pressure p by the increase in volume v is represented in the diagram by the area $ABCO$, or
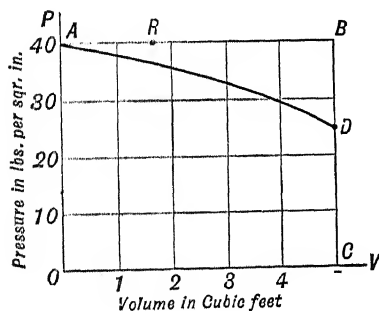
FIG. 644

the area between the line AB and the base line, measured in appropriate units. If pressures are measured in pounds per square foot, and volumes in cubic feet, the work pv will be given in foot-pounds. In the above case the work is $5 \times 40 \times 144 = 28,800$ ft.-lbs., and the area of each small rectangle in the diagram represents 1440 ft.-lbs. of work.

In case the pressure drops off as the volume increases, as indicated by the line AD in the diagram (Fig. 644), the work of expansion is represented by the area $ADCO$ between the curve AD and the base line OC. For imagine the increase in volume to be made by a succession of small steps, one of which is represented by v in figure 645. Then, if during the expansion v the pressure changes from p_1 to p_2, the work done will have a value between p_1v and p_2v, and will be greater than the shaded strip in diagram I and less
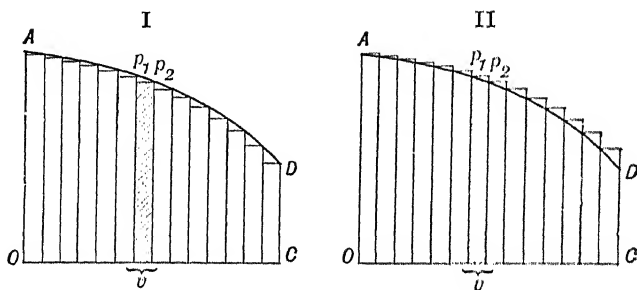
FIG. 645

than the corresponding strip in diagram II. The total work done in the expansion from A to D will be greater than the sum of all the little shaded strips in diagram I and still be less than those strips in diagram II. But the greater the number of strips in the expansion, the smaller will be the width of the strips and the more nearly will the sum of the areas in both cases approach the area $ADCO$ as a limit; that area, therefore, must represent the actual work in the expansion from A to D.

Hence in general *the work of an expansion represented by any line AD in the pressure-volume diagram is equal to the area under that line down to the base line, or line of zero pressure.*

3. Carnot's Cycle. In the year 1824 a French engineer, Sadi Carnot, then only 28 years of age, published a work of extraordinary originality on "the motive power of heat," in which he arrived at results of fundamental importance through the consideration of the properties of an ideal heat engine, in each complete

stroke of which the working substance was conceived to be put through a special series of four changes known as a Carnot's cycle, by which it was brought again to its original condition.

The working substance S is supposed to be enclosed in a cylinder having side walls and piston absolutely non-conducting, as indicated by heavy lines in the diagram, while the bottom is a perfect conductor of heat. Three stands are provided upon which the cylinder may be placed, a perfectly conducting hot stand at temperature T_2, a perfectly conducting cold stand at temperature T_1, and a perfectly non-conducting stand N.

Suppose the cylinder has been standing on the cold stand, and the working substance S, which may be supposed to be steam or air or any substance whatever, has come to the temperature T_1 and has
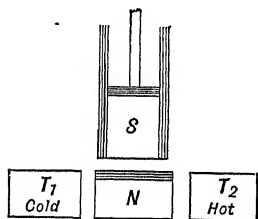
FIG. 646. Carnot's engine   FIG. 647
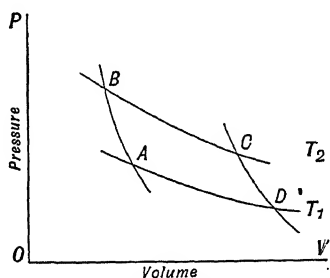
a pressure and volume represented by the point A on the diagram (Fig. 647).

1. The cylinder is transferred to the non-conducting stand N and the substance is compressed until its temperature rises to T_2 in consequence of the work expended in its compression, and comes to the volume and pressure represented by B. In this operation no heat can flow into or out of the substance; it is therefore called an *adiabatic* operation or change.

2. The cylinder is shifted to the hot stand and the substance allowed to expand. As it expands it does work and would be cooled except that heat freely flows in from the hot stand keeping its temperature constant at T_2. Let it expand in this way by some convenient amount until its volume and pressure may be represented by C on the diagram. The change from B to C is called *isothermal* because the temperature has remained constant.

3. Now place the cylinder again on the non-conducting stand, and let S expand still further. It will cool in consequence of the work done in the expansion since no heat is supplied, and may

be cooled in this way to the temperature T_1 of the cold stand. This change is *adiabatic* and represented by the line CD on the diagram.

4. Placing the cylinder on the cold stand the piston is pushed down compressing the substance from D to its original volume at A. During this operation the heat of compression is taken up by the cold stand as fast as it is developed and the temperature of the substance is thus kept constant at T_1. This change is *isothermal*, and while it is taking place heat flows out of the working substance into the cold stand.

This series of operations, in which the working substance is subjected to two isothermal and two adiabatic changes and brought again to its original pressure volume and temperature, is known as a Carnot's cycle; its peculiarity is that no heat transfer takes place except at two definite temperatures.
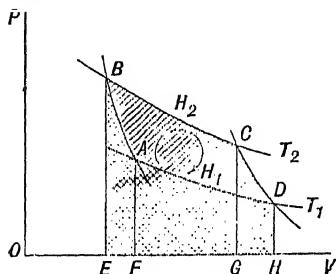
4. **Work in a Carnot's Cycle.** The work done by the working substance in expanding from B to C is represented by the area $EBCG$ (Fig. 648), and during the adiabatic expansion from C to D it does work represented by the area $GCDH$, while the work done *upon* the substance as it is compressed first from D to A and then from A to B is measured by the double crossed area $EBADH$, below the line DAB. The net amount of work obtained from the engine in a complete cycle is then represented by the excess of the work done *by* the substance in expanding, over the work done *upon* it in the compression, and is given by the area $ABCD$.

FIG. 648

During the isothermal expansion BC at the higher temperature T_2 heat is taken in from the hot stand, and during the isothermal compression DA heat is given out to the cold stand, while no heat at all is transferred during the two adiabatic changes. But the engine has done an amount of work during the cycle equal to the area $ABCD$, and since the working substance is in exactly the same state at the end as it was in the beginning, the energy expended in work cannot have come from the substance, but must have come from the energy supplied in the form of heat, otherwise the law of the conservation of energy would be violated. Experiment confirms this conclusion and we find that *the heat H_2 taken in by the substance in expanding from B to C is more than the heat H_1 given out in the compression from D to A, the difference between the two being mechanically equivalent to the work of the cycle $ABCD$,*

so that if work and heat are measured in the same units, which can be done since both are forms of energy, we have

$$W = H_2 - H_1$$

where W is the work represented by the area $ABCD$.

5. **First Law of Thermodynamics.** The conclusion just reached is based on what is known as *the first law of thermodynamics*, which is simply the law of the conservation of energy as applied to the relation of heat to work. It asserts that *wherever work is obtained by any heat process an equivalent amount of heat disappears, and vice versa.*

6. **Efficiency of an Engine.** The efficiency of a heat engine is the ratio of the work which the engine does to the energy that has to be supplied to it from the hot stand or boiler. In the case just discussed the efficiency E may be expressed by the formulas

$$E = \frac{W}{H_2} \quad \text{or,} \quad E = \frac{H_2 - H_1}{H_2}.$$

7. **Reversible and Irreversible Operations.** Some operations are in their energy relations reversible and some are irreversible. For instance when a weight is raised energy is expended, and when it is lowered again an equal amount is given back. This is a reversible operation; but when work is done against friction energy is transformed into heat, and when the motion is reversed energy is not recovered but still more is spent in heat, and the operation is not reversible. So the conduction of heat from hotter to colder bodies is an irreversible operation, while the adiabatic heating or cooling of a substance, as its volume is changed without any conduction of heat taking place, is a reversible change.

A reversible engine is one in which the operations are all reversible. If an engine working in a Carnot's cycle were to be made so that during the isothermal expansion in which heat is taken in from the hot source there were absolutely no difference of temperature between the working substance and the source, and if there were a corresponding transfer of heat with no difference of temperature during the isothermal compression while heat is given out, the operations would then all be reversible.

If such an engine were driven backward the working substance would be put through the cycle of operations represented in figure 647 in the reverse order. Expanding from A to D it would take in the same quantity of heat H_1 at the cold temperature T_1 as it had given out at that temperature when direct acting; and in the

compression from C to B it would give out the same quantity of heat H_2 to the hot body as it had taken in when direct acting. And in this reversed action the heat given out would be more than that taken in, by the work of the cycle W, which would in this case have to be supplied from outside, for more work would be done upon the working substance in the compression than would be done by it in the expansion.

It is evident that no actual engine is ever reversible, nevertheless the results obtained by considering the properties of such engines are of very great importance.

8. Carnot's Theorem. *No engine can be more efficient than a reversible engine working between the same limits of temperature.* For, if possible, let some engine E be more efficient than the reversible engine R working between the same limits of temperature T_1 and T_2, and suppose the strokes of the engines are so adjusted that the work done per cycle is the same by one as by the other. Then let the more efficient engine E be coupled to the reversible engine R so as to drive it backward or in the reversed direction. This it will be able to do for the work required to drive the reversible engine backward is exactly what it would have performed if direct acting. In every cycle the engine E takes in heat H at the upper temperature and gives out heat h at the lower temperature and the work W which it does is equal to $H - h$ as we have seen (§ 4). The reversible engine, on the other hand, takes in heat h' at the lower temperature and gives out heat H' at the higher temperature and in this case also $H' - h' = W$; and so, since W is the same for one as for the other, we have $H - h = H' - h'$. But by hypothesis the efficiency of E is greater than the efficiency of R, that is,
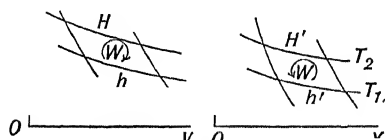
$$\frac{W}{H} > \frac{W}{H'}, \text{ therefore } H < H' \text{ and } h < h'.$$

FIG. 649

That is, the heat H' returned to the boiler is more than the heat H taken from it, and the heat h' taken out of the cold body is more than the heat h which is given to it. There results, therefore, a steady transfer of heat from the cold to the hot body, the cold body growing colder and the hot body hotter through the agency of the combined engines working continuously without any outside assistance. This result is believed to be impossible, for it contradicts

all experience. This conviction when formulated is called *the second law of thermodynamics* and may be stated thus: *It is impossible by any continuous self-sustaining process for heat to be transferred from a colder to a hotter body.*

We conclude then that *no engine can be more efficient than a reversible engine working between the same limits of temperature,* and consequently, *all reversible engines working between the same limits of temperature are equally efficient,* whatever working substance is used in the engine, whether air, steam, gas, or substance of any sort.

9. Absolute Temperature Scale. Let us suppose that the diagrams in figure 650 are for two different gases or vapors between the same limits of temperature T_1 and T_2, then according to the theorem just given a reversible engine working with one substance in the cycle $ABCD$ will be exactly as efficient as another reversible engine using the other substance and working around the cycle $A'B'C'D'$. Consequently

$$\frac{W}{H} = \frac{W'}{H'}$$

where W and W' are the work areas of the two cycles and H and H' are the quantities of heat taken in at the temperature T_2.

Since the ratio $\frac{W}{H}$ depends only on the temperatures T_1 and T_2 and not at all on the kind of substance used, Lord Kelvin proposed making it the basis of a truly absolute scale of temperature which should be quite independent of the individual peculiarities of different substances. This scale may be easily understood by the
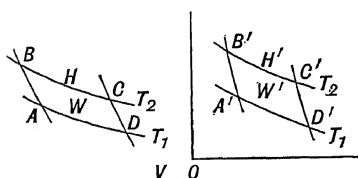
FIG. 650
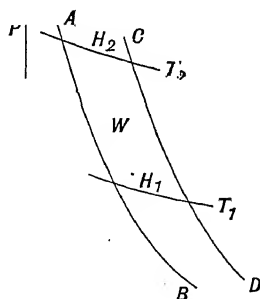
FIG. 651

aid of figure 651, which is the pressure volume diagram of some substance, for which AB and CD are two adiabatic curves. Suppose T_1 and T_2 are two standard temperatures used to fix the size of the degree of the scale. For instance T_2 may be the standard boil-

ing temperature of water, and T_1 the freezing point of water, and we may decide upon a scale in which, as in the Centigrade scale, there shall be 100° between the two temperatures. Then if isothermal lines for the substance are drawn between T_1 and T_2 so as to divide the whole area W into 100 equal parts, these will correspond to successive degrees of temperature of the Kelvin scale, and the area between any two isothermal lines which are 1° apart will be one one-hundredth of W, or a.

But in passing from the adiabatic line AB to the adiabatic CD at the temperature T_1 a quantity of heat energy H_1 is absorbed by the substance which is less than H_2, the amount absorbed at T_2, a temperature 100° higher, by an amount equal to W, or $100a$, and at 1° below T_1 the heat taken in in passing from one adiabatic to the other will be less than H_1 by a; at 2° below T_1 it will be less by $2a$; and so on. But if we take as many degrees below T_1 as a is contained in H_1, we shall reach a temperature at which no heat at all would be taken in or given out in the isothermal change from one adiabatic to the other. Now, if this temperature is taken as the lower one in a Carnot's cycle, and any other temperature be taken as the higher one, a reversible engine working in the cycle will have unit efficiency, that is, all the heat taken in at the upper temperature will be transformed into mechanical work, and none will be given out at the lower temperature.

No higher efficiency than this can be possible and the lower temperature thus defined is taken as the absolute zero.

Experiment shows that H_1 is nearly 273 times a, so that the freezing point of water is about 273° above the absolute zero of the Kelvin scale, or as we may write it 273° K., and the boiling point being 100° higher will be 373° K.

If 180° had been taken between the freezing and boiling points of water, as in the ordinary Fahrenheit scale, the freezing temperature of water would have been found 491° and the boiling point 671° above the absolute zero.

It is interesting to note that temperatures measured on the absolute scale of Lord Kelvin agree very closely with temperatures measured from the so-called absolute zero of the air thermometer (§ 394); but whereas the zero of the air thermometer was based on the behavior of a particular class of substances, the gases, the zero of the Kelvin scale is independent of the properties of any particular substance.

10. Efficiency of a Heat Engine on Kelvin's Scale. From the above explanation of the Kelvin scale of temperature it will be seen that *the heat taken in by a substance in passing from one given*

*adiabatic to another given adiabatic at any temperature is propor-
tional to the number of degrees on the Kelvin scale which that tem-
perature is above the absolute zero; that is,*

$$\frac{H_1}{T_1} \quad \frac{H_2}{T_2} = a$$

or, $H_1 = aT_1$ and $H_2 = aT_2.$

But we have seen (§ 6) that the efficiency E of a reversible en-
gine may be expressed in terms of the heat H_2 taken in at the higher
temperature and the heat H_1 given out at the lower temperature
by the relation

$$E = \frac{H_2 - H_1}{H_2}$$

So that substituting the values just given for H_1 and H_2 we have

Efficiency of a reversible engine acting $\left.\begin{array}{c} \\ \end{array}\right\}$ $= \dfrac{T_2 - T_1}{T_2}$
between the temperatures T_1 and T_2

where the temperatures are measured on the Kelvin absolute scale.

But no engine is more efficient than a reversible engine working
between the same limits of temperature; hence the expression
gives an upper limit to the efficiency of any heat engine whatever.
It will be noted that to secure high efficiency in a steam engine there
must be great difference in temperature between the steam as it
enters and as it escapes, and the high efficiency of gas engines is
due in part to the high initial temperature of the exploding mixture.

PROOF OF NEWTON'S WAVE FORMULA

The following proof of Newton's law that in case of a compressional wave $V = \sqrt{\dfrac{E}{d}}$ is due to the distinguished English engineer and physicist, W. M. Rankine. Let a sound wave be moving forward in a medium in the direction and with velocity indicated by the arrow V, and let A and B be two parallel planes, perpendicular to V, which move forward with the same velocity as the wave, and therefore are always the same distance apart.

Let u represent the velocity of a particle due to its oscillation relative to the medium, p the pressure in the medium, d the density of the medium, v the volume of unit mass, $v = \dfrac{1}{d}$;

and let the values of these quantities at the plane A be represented by u_1, p_1, d_1, v_1 and at plane B by u_2, p_2, d_2, v_2.

Now suppose for definiteness that the plane A is at the point of maximum compression of the wave, then u_1 will be maximum and in a forward direction

FIG. 652

as indicated in the figure, p_1 and d_1 will both have their maximum values, and these values remain constant as the plane moves on, for it moves *with the wave*. So also the conditions at plane B, at some other point in the wave, though different from those at A remain constant as the plane moves on.
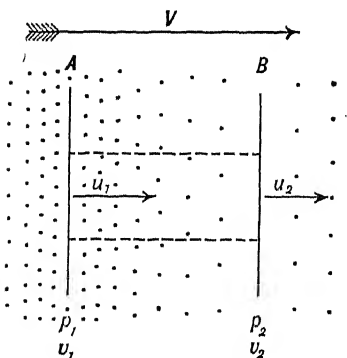
To an observer moving with the planes A and B the medium will be seen to stream through the planes from right to left. At B its velocity relative to the plane will be $V - u_2$ and this then will be the number of cubic centimeters of the medium that pass through each square centimeter of B in one second. Let M be the *mass* of this volume, and we have

$$Mv_2 = V - u_2. \tag{1}$$

The mass M passing B per second must be the same as the mass passing A in the same time otherwise there will be a change in the amount of matter between A and B as the wave moves on, therefore

$$Mv_1 = V - u_1. \tag{2}$$

At a point where the velocity of the particles is zero the medium has its normal density d, and normal volume per unit mass v, hence at that point

$$Mv = V. \tag{3}$$

Now consider a column of the medium reaching from A to B and having 1 sq. cm. cross section, the mass M entering this region through B in one second has momentum Mu_2, and the same mass leaving at A has momentum Mu_1 which we have supposed greater. The column thus experiences a *loss of momentum per sec.* equal to $Mu_1 - Mu_2$. But the conditions in the column *do not change* as the wave moves along, so the loss must be balanced by an equal gain in momentum. This is supplied by the difference between the pressures p_1 and p_2 on the two ends of the column, p_1 urging it toward the right and p_2 toward the left. The column must, therefore, by Newton's second law of motion, experience a gain in momentum per second from left to right, equal to $p_1 - p_2$; and this momentum must be equal to that which the column loses in the same time.

We have therefore $\qquad M(u_1 - \qquad \qquad \qquad$ (4)

But subtracting (2) from (1)

$$(u_1 - u_2) = M(v_2 - v_1)$$

whence $\qquad M^2(v_2 - v_1) = p_1 - p_2.$

And by equation (3)

$$\frac{V^2}{v^2} = \frac{p_1 - p_2}{v_2 - v_1} \quad \text{or} \quad V^2 \quad v \left| v \frac{p_1 - p_2}{v_2 - v} \right.$$

But by §§ 247 and 249, $\quad E = v \dfrac{p_1 - p_2}{v_2 - v_1}, \quad$ therefore, $\quad V^2 = vE.$

And since $v = \dfrac{1}{d}$ we have finally $V = \sqrt{\dfrac{E}{d}}.$

APPENDIX III

THE BOHR THEORY

1. Equilibrium of Revolving Electron in Hydrogen Atom. Bohr selected the hydrogen atom for investigation because of its comparative simplicity. He made the natural assumption that Coulomb's law of attraction existed between the nucleus and the electron, and that the electron was held in a definite orbit, and prevented from falling into the nucleus, by the centrifugal force produced by its velocity of revolution in its orbit. This idea was not new, but it plays an important part in the Bohr atom. Coulomb's law of attraction between two opposite electric charges will be recalled to be of the form (see § 539)

$$F = \frac{qq'}{Kr^2}. \tag{1}$$

In the case of an atomic nucleus of charge E and an electron of charge e, this equation takes the form

$$F = \frac{Ee}{a^2} \tag{2}$$

where the dielectric constant K is unity, and the distance between the atomic nucleus and the electron is called a to conform with usual terminology. Now if the electron is assumed to be revolving in a circular orbit, it exerts a centrifugal force equal to $\frac{mv^2}{a}$ (§ 116). For the circular motion to be maintained, this outward centrifugal force must just balance the inward pull of the nucleus upon the electron, or

$$\frac{mv^2}{a} \qquad \frac{Ee}{a} \tag{3}$$

An exact expression should take account of a small circular motion of the nucleus itself which is present also but since the nucleus has over 1800 times the mass of the electron, this is very small and has been neglected in equation (3).

2. Determination of the Energy of the Bohr Atom. Next, how is the *energy* of such a system to be measured? This consists of two parts, namely, the potential energy of the system, due to the force of the electron upon the nucleus, and the kinetic energy, due to the velocity

of the electron in its orbit. If the electron approaches the nucleus, its potential energy *decreases;* on the other hand it is *increased* by pulling the electron away from the nucleus, just as the potential energy of a weight increases as it is lifted above the earth. The potential energy of the electron with respect to the nucleus can be shown to be

$$\frac{Ee}{a} \tag{4}$$

where a is its distance from the nucleus. This expression is derived by a method exactly like that used in § 591 in the discussion of potential except that the electron charge e replaces the unit charge so that the expression on page 409 becomes $V = \frac{qq'}{r}$ in place of $\frac{q}{r}$, where as before $q = E$, $q' = e$, and $r = a$. The negative sign means that, since the electron is attracted instead of repelled, work is done *by* the electron (in moving it toward the nucleus) instead of *upon* it as would be true in a case of repulsion.

The kinetic energy of the electron is evidently determined by its orbital velocity v, and is equal to $\frac{1}{2}mv^2$ where m is the mass of the electron. From equation (3) $mv^2 = \frac{Ee}{a}$, or $\frac{1}{2}mv^2 = \frac{1}{2}\frac{Ee}{a}$. Thus the kinetic energy is seen to be equal to $\frac{1}{2}\frac{Ee}{a}$. $\tag{5}$

Since the *total* energy W of the system is equal to the sum of its potential and kinetic energies, W is obtained by simply adding equations (4) and (5) which gives $W = -\frac{1}{2}\frac{Ee}{a}$ $\tag{6}$

where a is the distance between the electron and the nucleus. The *total* energy is thus greater the farther the electron is from the nucleus according to the law of equation (6) and evidently becomes zero when the electron is completely removed from the atom. This is, however, entirely a matter of convenience in using equation (6). The zero value of energy might be made anything desired by adding a term to (6) representing a given constant amount of energy.

3. **Energy Changes During Absorption and Emission of Radiation.** It is only energy *changes* which are actually observed by experiment. If, for instance, the electron is moved from an orbit of radius a_1 to one of larger radius a_2, the energy change is equal to the final energy minus the original energy which is equal to

$$\Delta W = \frac{1}{2}\frac{Ee}{a_1} - \frac{1}{2}\frac{Ee}{a_2} \tag{7}$$

where ΔW is the energy change. If the zero point of energy were altered by adding or subtracting a certain constant quantity of energy to W (equation 6) this constant quantity is seen to cancel out in expressions for energy change like that of equation (7).

Thus for *convenience* the energy of the electron is taken as zero when removed from the atom, and thought of as decreasing step by step to greater and greater negative values as the distance a becomes less and less.

If the energy change ΔW in equation (7) is placed equal to a quantum $h\nu$ this equation becomes

$$h\nu = \frac{1}{2}\frac{Ee}{a_1} \cdot \frac{1}{2}\frac{Ee}{a_2} \tag{8}$$

This equation is seen to have exactly the same form as equation (7), (page 731) which will be written again

$$h\nu = \frac{Rhc}{n_1^2} - \frac{Rhc}{n_2^2}. \tag{9}$$

Thus it is seen by comparing these two equations that a quantum of energy $h\nu$ may be thought of as emitted by a change of energy of the atom produced by the electron shifting from an orbit of radius a_2 to the smaller radius a_1, and vice versa for the absorption of the quantum. Furthermore a_1, and a_2 must have values which make the two right-hand terms of equation (8) exactly equal to those of equation (9).

4. Determination of the Electron Orbits Corresponding to the Energy States. Bohr then found by what law these distances must be determined to make equation (8) identical with equation (9). He found that if he assumed that

$$mva = \frac{nh}{2\pi} \tag{10}$$

the series of values of a found by giving n its series of integral values was the required series of distances of the electron from the nucleus to make equation (8) and (9) give identical series of energy differences, or identical series of quanta. This equation says that the stationary states of the Bohr atom are such that the *angular momentum* of the electron about the nucleus is a whole multiple of $\frac{h}{2\pi}$ where h is Planck's constant. The quantity $mva = ma^2\omega$ (since $v = a\omega$) is known as the angular momentum of the electron about the nucleus, where ω is the angular velocity of the electron (§ 137) and ma^2 corresponds to moment of inertia I (§ 144).

By eliminating v between equations (3) and (10) and solving the resulting equation for a, it is found that

$$a = \frac{n^2 h^2}{4\pi^2 m E e}.$$ (11)

If this value of a is substituted in equation (8) it becomes

$$h\nu = \frac{2\pi^2 m E^2 e^2}{\quad} - \frac{2\pi^2 m E^2 e^2}{\quad}$$ (12)

where $n = n_1$ when $a = a_1$ and $n = n_2$ when $a = a_2$.

5. Evaluation of the Rydberg Constant. If the numerators and denominators of the right-hand numbers of equation (12) are each multiplied and divided by ch this equation becomes

$$h\nu = \frac{2\pi^2 m E^2 e^2 ch}{ch^3}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) = Rch\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$ (13)

The nuclear charge of the hydrogen atom consists of one positive unit equal in magnitude but opposite in sign to that of the electron. Thus E in equations (2) to (13) can be replaced by e, so that

$$R = \frac{2\pi^2 m e^4}{ch^3}.$$ (14)

This is the correct expression for the Rydberg constant. For other atoms E is greater than unity. For helium $E = 2$, and lithium $E = 3$, etc., so that the constant Rch of equation (9) becomes $4Rch$, $9Rch$, etc., R always being given by the expression (14).

6. Relation Between Orbital Frequencies and Radiation Frequencies. The number of revolutions per second f of an electron about the nucleus is equal to its velocity divided by the circumference of its orbit $= \dfrac{v}{2\pi a}$, that is,

$$v = 2\pi a f.$$ (15)

If this value of v is substituted in equation (10) the frequency of revolution is found to be

$$f = \frac{nh}{4\pi^2 m a^2}.$$ (16)

Substituting in this the value of a from equation (11) and making use of equation (14) and replacing E by e, since for hydrogen $E = e$,

$$f = \frac{2Rc}{n^3}.$$ (17)

Comparing this with equation (6) § 1023, for the vibration frequencies of spectral lines of the hydrogen atom it appears that the orbital frequency and the frequency of radiation are by no means the same and that they depend on two very different laws. This is contrary to the electromagnetic theory, which is abandoned by Bohr as applied to the absorption and emission of radiation by atoms. It is very interesting, however, that for large values of n, when n_2 is one unit greater than n_1 such as 101 and 100, the revolution frequencies and the quantum frequencies almost exactly coincide, exact coincidence taking place when n_2 and n_1 differ by one unit and are indefinitely large. This was made the basis of Bohr's famous *principle of correspondence* between electron vibration frequencies and spectrum line frequencies, a further discussion of which can be found in more advanced works on this subject.